Assessing Investigative Skill Development in Inquiry-Based and Traditional College Science Laboratory Courses

Jerry P. Suits
University of Northern Colorado

A laboratory practical examination was used to compare the investigative skills developed in two different types of general-chemistry laboratory courses. Science and engineering majors (SEM) in the control group used a traditional verification approach (SEM-Ctrl), whereas those in the treatment group learned from an innovative, inquiry-based approach (SEM-Trt). A scoring rubric was developed from their examination sheets to assess six component investigative skills. Results indicated that SEM students in the SEM-Trt group scored significantly higher than those in SEM-Ctrl for all six skills. Furthermore, nursing and applied science majors (NonSEM) in the inquiry-based group (NonSEM-Trt) wrote significantly better discussions than did SEM students in SEM-Ctrl group. Overall, competency at the mid-range level of laboratory skills was attained by most SEM-Trt students (72.5%) but by only 30.5% of SEM-Ctrl and 28.6% of NonSEM-Trt students. Apparently, during the semester students in the SEM-Trt group were able to use the inquiry-based method effectively to combine chemical tasks with writing tasks and postlaboratory discussions. One implication of this study for science instructors is that practical examinations can provide useful feedback regarding the quality of the laboratory experience. Another implication is that this study provides evidence for the use of the innovative inquiry-based laboratory approach to support student learning of high-level investigative skills. However, students' requisite background knowledge must match the level of these skills.

"Science as inquiry" is the instructional goal advocated by the *National Science Education Standards* (National Research Council, 1996). Educators have recognized that implementing this goal in schools requires strategies to increase instructional effectiveness and to align assessment strategies with standards (Bybee, Ferrini-Mundy, & Loucks-Horsley, 1997). In chemistry, the laboratory is a logical instructional environment in which this goal could be met. However, some chemical educators have questioned whether the laboratory experience actually makes a unique and essential contribution to chemistry instruction (Horton, 1928; Lagowski, 1989; Lock, 1990). Most laboratory instructors tend to believe that laboratory instruction improves manipulative skills, observational skills, and a number of higher order cognitive skills that reflect an understanding of the scientific investigative method (DeMeo, 1997; Meester & Maskill, 1995; Shiland, 1999). In recent years, however, the assumptions underlying this belief have been questioned (Hilosky, Sutman, & Schmuckler, 1998; Lagowski, 1989; Laws, 1996; Ryder & Leach, 1999), for example:

- How are students supposed to decide what constitutes "good data"?
- How can students use data as evidence to assess the validity of scientific models?
- Should students be expected to think like research scientists, which would require development of an integrated set of complex, interrelated process skills?

These concerns need to be considered whenever science instructors intend to design or modify, implement, and subsequently, evaluate a laboratory program.

The extent to which laboratory work can help students develop their investigative skills depends upon which instructional method (approach) is selected. For many years the *verification,* or cookbook, approach has been used extensively, despite the fact that it emphasizes laboratory techniques while de-emphasizing investigative skill development (Abraham et al. 1997; Beasley, 1991; Hilosky et al. 1998; Lagowski, 1989; Laws, 1996). In this approach students are given a predetermined laboratory outcome (expected results), which they are to verify by following a step-by-step procedure, by filling in blanks on data sheets, and by answering a set of questions that usually query knowledge of facts or calculation algorithms. This deductive approach encourages only "rule-governed behavior" (Beasley, 1991); however, it remains popular among laboratory instructors because it consumes minimal resources in terms of time, space, equipment, and/or personnel (Lagowski, 1990). Conversely, the *inquiry* approach maximizes demands for these resources in addition to the expectation that students discover and develop their own investigative skills (Huber & Moore, 2001). These skills are then used by students to write, execute, and analyze their own procedures and to discover inductively underlying chemical principles and undisclosed laboratory outcomes (DiBiase & Wagner, 2002). Fortunately, chemists and educators have generated two hybrid approaches:

- A *guided inquiry* approach that uses an inductive sequence (just like inquiry), but it provides some guidance to students, who are given a procedure and a predetermined outcome, which they are expected to discover (Domin, 1999).
- A *problem-based* approach that uses a deductive sequence to provide students the theoretical basis with which to generate their own laboratory procedure and then to investigate a predetermined outcome (Domin, 1999; Goodman & Bean, 1983; Laws, 1996).

Instructors can select whichever approach matches the level of laboratory competence (Beasley, 1991; Meester & Maskill, 1995) they expect their students to develop. Each of these four instructional approaches can be placed on a continuum of laboratory competence that ranges from the verification method (Level I or II), which can develop primarily manipulative skills (i.e., technique), to the other three methods that demand higher levels of investigative skills (Meester & Maskill, 1995). Guided inquiry can be classified at Level III (Beasley, 1991), because although it gives students a stepwise procedure to follow, they are expected to gather and analyze evidence to generate their own conclusions and to discover underlying chemical principles. Meanwhile, problem-based instruction satisfies the criteria for Level IV (Beasley, 1991), because it demands that students generate their own procedure based upon cues given in the statement of the problem to be solved (DeMeo, 1997; Meester & Maskill, 1995). The *open inquiry method* combines the separate demands imposed by guided inquiry and problem-based forms of instruction (Level IV+), because students must use an inductive approach in which they generate their own procedure and then discover underlying chemical principles.

After science instructors have selected and implemented an instructional approach based upon level of student competence desired, they probably want to gather feedback on whether their students have met the criteria defined at that particular level. Although written laboratory examinations may be convenient to prepare and grade, this assessment format does not engage students in the three stages of scientific investigation: first, in the prelaboratory stage they must plan how to do the chemical task, in the experimental stage they actually perform the chemical task, and then in the post-laboratory stage they analyze their results. Conversely, a practical laboratory examination can be designed to assess laboratory competence during all three stages of scientific investigations (Goh, Toh, & Chia, 1989; Hilosky et al., 1998). Silberman and his colleagues (1987) designed practical examinations that assessed student performance on six categories of laboratory skills.

They concluded that a practical examination has a twofold beneficial effect, because it encourages (a) laboratory instructors to examine their goals and objectives, and (b) students to become aware of their laboratory skills developed during the semester. For the second case, without a final laboratory examination, students tend to forget about the laboratory activities as soon as they are completed. However, if students are forewarned that the laboratory program culminates with an examination, then they tend to keep better laboratory notebooks and to review all of the experiments performed during the semester.

When students are taught with an instructional approach that imposes a higher level of laboratory competence, that is, Level III or IV (Beasley, 1991), they must develop their writing skills in order to match the higher level of demand. Goodman and Bean (1983) used a problem-based approach, Level III, to emphasize writing skills as an attempt to simulate actual research conditions and, thus, promote the growth of problem-solving and analytical skills. They found that the real challenges for students were devising, executing, and later, explaining in writing the procedure for conducting a systematic scientific investigation. They concluded that these writing tasks compelled students to generate their own ideas in response to a problem-solving situation and to organize and clarify their ideas.

Keys (2000) also found that the act of writing can engage students in high levels of scientific thinking. If an instructional approach combines a writing task with a chemistry-based task, then students can practice integrating their component investigative skills into a coherent form. Thus, a laboratory practical examination can be used to assess the effectiveness of particular instruction by determining whether or not student practice of investigative skills allows them to integrate these skills into a coherent set representing a higher level of laboratory competence.

Purpose

This paper describes how a laboratory practical examination can assess scientific investigative skills acquired by students taught by two instructional methods (control and treatment, see Table 1) that emphasize different levels of laboratory competence. Three student groups (independent variable) participated in this study: science and engineering majors (SEM) in control (SEM-Ctrl) and treatment (SEM-Trt) groups, and non-SEM majors in the second treatment group (nonSEM-Trt). The dependent variable was their performance regarding six scientific investigative skills (see appendix). The specific research questions were as follows:

1. What is the effect of laboratory instructional method upon the acquisition of six scientific investigative skills (quantitative measure)?
   a. pairwise t-test comparison of SEM-Ctrl and SEM-Trt groups.
   b. pairwise t-test comparison of SEM-Trt and NonSEM-Trt groups.
   c. pairwise t-test comparison of SEM-Ctrl and NonSEM-Trt groups.
2. How does laboratory instructional method affect the number of students who acquire a coherent set of investigative skills at the "midrange competent" and the "fully competent" levels (quantitative measure)?
3. What are student perceptions of their thinking skills used during a practical examination (qualitative method)?

Methodology

Description of Sample

This study was conducted at a midsized state university in the South (U.S.) that has an open admissions policy. The average ACT-Mathematics score for SEM students enrolled in introductory-level chemistry courses ranged from 20 to 22. The need for this study arose after the chemistry department decided to change its introductory-level chemistry curricula from a traditional structure to one that imposes prerequisites for laboratory courses and allows instructional innovation. Prior to this change, students completed two semesters of general chemistry. Each four-credit course included lecture (80% of grade) and laboratory (20%) components. The new curricula consisted of a first-semester, three-credit lecture course followed by a second-semester, three-credit lecture course coupled with a co-requisite two-credit laboratory course. The laboratory course featured a 1-hour prelaboratory lecture, a 3-hour laboratory, and 1-hour postlaboratory discussion. Students wrote a laboratory report for each laboratory activity completed. Differences between the two laboratory approaches are summarized in Table 1.

In this study, two groups were both composed of SEM students enrolled in second-semester general chemistry; that is, the SEM control group (59 students in SEM-Ctrl, fall semester) and the SEM treatment group (51 students in SEM-Trt, spring). A second treatment group was drawn from a different population (NonSEM), composed of nursing and applied science majors (42 students in NonSEM-Trt). During the fall semester, SEM-Ctrl students performed verification (cookbook) activities, Level II. The focus was on students doing experiments (one 3-hour laboratory period weekly for two semesters) but less emphasis was placed on students' prelaboratory preparation (a prelaboratory quiz and brief lecture) and no emphasis on postlaboratory analysis. During the spring semester, students in the two treatment groups participated in an instructional method that featured all three stages of laboratory investigation (Level III), that is, prelaboratory preparation, experimental work, and postlaboratory analysis (see Table 1).

This study assumed that the SEM-Ctrl and SEM-Trt groups were equivalent, because both groups were drawn from the same SEM population taking general chemistry at one university. A check on this assumption based upon specific background statistics of these two groups is shown in Table 2.

The Laboratory Assessment Method

Laboratory practical examinations, called Laboratory Challenge Experiments, were used in this study to assess the laboratory performances of students as they wrote their own procedure and investigated chemical phenomena (Level IV). The faculty developed these Laboratory Challenge Experiments as part of the regular instructional program. These examinations, used in the fall and spring, posed similar quantitative problems (Table 1) that compared the acquired skills of students in the control and the two treatment groups. All groups were informed at the start of the semester that they would take a laboratory practical examination at the close of the semester based loosely upon their previous laboratory work. However, no details about the examination were provided to students during either the control or treatment semester.

Table 1
Comparison of Instructional Methods for Control (Fall) and Treatment Groups (Spring)

Stage or Activity	Traditional Instructional Method (SEM-Ctrl)	Inquiry-Based Instructional Method (SEM-Trt & NonSEM-Trt)
Instructional Approach (Level)	Verification or Cookbook (Level II).	Problem-based (Level III).
Pre-laboratory preparation	Study for pre-laboratory quiz.	Attend pre-laboratory lecture, write objective and procedure in laboratory notebook, and anticipate "expected results."
Types of experiments	Traditional wet labs.	Traditional wet labs plus computer-interfaced experiments.
Experimental Work	Fill in the blanks; science and engineering majors worked in pairs.	Fill in the blanks; key data into a computer to obtain class mean & standard deviation; science and engineering majors worked individually on most experiments.
Post-laboratory analysis	None (turn in report sheet(s) before leaving the lab).	Oral discussion followed by a written discussion of the results obtained, error analysis, and significance of the experiment.
Written Examinations	Memorization and computation; (format—multiple choice only).	Statistics, graph construction and interpretation, some memorization, computation, and essay questions that required solving multi-step problems (65% multiple-choice and 35% essay).
Practical Exam: Laboratory Challenge Experiment	Problem: In the laboratory room, you will be given two white powders. One is composed of an organic acid, while the other is a metal chloride (the metal is a divalent cation). Your mission is to use quantitative procedure(s) to identify the organic acid from its properties, and to use qualitative test(s) to identify the metal chloride.	Problem: In the laboratory room, you will be given a white powder. It is a mixture of an organic acid and an ionic compound. Your mission is to use quantitative procedure(s) to determine the percent acid in the impure sample.

Table 2
Comparison of theTwo Groups of SEM Majors on Three Different Parameters That Can Affect Science Achievement

Parameter	SEM-Ctrl Mean (*SD*)	SEM-Trt Mean (*SD*)
ACT-Mathematics	20.2 (4.4)	21.5 (4.7)
ACT-Composite	20.7 (3.9)	21.8 (4.1)
Prior Chemistry Knowledge[a] (GPA)	2.04 (1.08)	2.16 (1.23)

[a]Prior chemistry knowledge was determined from students' grades in first semester general chemistry— the prerequisite chemistry course for both SEM groups.

During a Laboratory Challenge Experiment, students in all three groups participated in the three stages of empirical scientific methodology.

First stage. Students were given about 30 minutes in the lecture room to read the stated problem, write their own procedure, and list any materials/equipment needed. The empirical problem was printed near the top of a legal-sized "challenge worksheet" that also included sections for procedure, observations/calculations, and discussion. During this writing task, students were allowed to refer to their laboratory lecture notes and previously graded laboratory reports. Thus, most devised a procedure based upon a modification of an activity they had completed earlier in the semester. Next, they turned in their challenge worksheets to their teaching assistant (TA) and took a 10-minute break, during which they were allowed to intermingle and talk freely in the hallway. Meanwhile, the TAs drew boundary lines around their written procedures and materials sections. This break gave students the opportunity to exchange information, that is, discuss their procedures, and predictions. This prelaboratory stage prepared them for their laboratory work.

Second stage. Students entered the laboratory room where a "silence rule" was enforced over the 2 hours of laboratory activity. SEM students completed their laboratory activity and gathered their own data, while students in the NonSEM sections worked quietly together in pairs. The room was equipped with all anticipated chemicals, glassware, and equipment needed for the student activity.

Third stage. Students finished writing their laboratory reports (data, revised procedure, and discussion) and turned them in before leaving the laboratory room. They were graded on their written plan/protocol, laboratory work (technique), results (accurate versus inaccurate data), and their written discussion.

For research purposes, the author and an undergraduate research assistant (RA) constructed a scoring rubric to classify students at various levels of laboratory competence (see appendix). The RA was a graduating senior with a good background in both chemistry and science education courses. The correlation between these two raters on total points (i.e., the sum of competence points for all six component skills listed in the appendix) was moderately high, $r = 0.86$, producing an interrater reliability of 74% agreement. The RA had no knowledge of the nature of this research project regarding the three groups (i.e., SEM- and NonSEM-treatment groups, and control-SEM group).

Quantitative Research Methods

The data were subjected to two types of analysis: quantitative methods and a qualitative method. Two quantitative methods compared student performance within the three groups. For Research Question 1, a t-test statistic ($p < 0.05$) was used in a pairwise comparison of group means on each of six component investigative skills. In addition, number of lines of discussion, regardless of line length or letter size, written by each student were counted, and the mean for each group was calculated. For Research Question 2, *laboratory competence level* was defined as follows: To be considered competent, a student needed to meet the criterion level on at least five of the six skills. Two levels were established: (a) a midrange competent level criterion was set at or above Level 2 for each skill, and (b) a fully competent level criterion was set at Level 3 (see appendix). The percent of students meeting criterion within each group was then calculated.

Qualitative Research Method

For Research Question 3, a qualitative method was used to evaluate written responses of treatment-group SEM students (SEM-Trt) on an essay question, which was part of the written final laboratory examination:

> Employers want college graduates who can think rather than recite memorized answers... Did any laboratory experiment or challenge that you did this semester meet these criteria? ____ If yes, what experiment? If no, why not?

No qualitative evaluation was possible for the control group (SEM-Ctrl) because they completed a different written final laboratory examination that included only multiple-choice items.

Results

Component Investigative Skills

For Research Question 1a, the pairwise t-test showed that students enrolled in the SEM-Trt treatment group scored significantly higher ($p < .05$) than did their cohorts in the control group (SEM-Ctrl) on all six component investigative skills (see Figure 1). Regarding specific skills, given a written statement of the problem, the SEM-Trt students were better able to plan and describe a procedure to use, make and record observations during the experiment, collect data, and then calculate and properly record their results. In addition, they wrote better (in the judgment of the two raters) and longer discussions, in which they first compared observed results with their expected results, and then used their observed results as evidence to verify whether the experimental objective was accomplished. In discussing these two factors, SEM-Trt students ($M = 6.02$ lines, $SD = 2.68$) wrote significantly longer discussions than did their SEM-Ctrl cohorts ($M = 3.17$ lines, $SD = 2.72$). Although both groups were drawn from the same SEM population from the same university, it is possible that there were group differences in ability, interest, or background that might have produced any of the differences in skill development found in this study.

For the results of Research Question 1b, pairwise t-tests were performed on the two treatment groups, that is, SEM majors in the SEM-Trt group and nursing and applied science majors in the NonSEM-Trt group. Both groups completed the same experiments, but they differed in that the SEM-Trt students were enrolled in a different lecture course that demanded more mathematical and scientific rigor to understand the scientific content of the course. The results showed that the SEM-Trt group scored significantly higher ($p < .05$)

than did the NonSEM-Trt group on five of six skill categories. The exception was the investigative skill, in which procedural steps were written during the planning stage (see appendix). There was no significant difference on this particular skill.

For Research Question 1c, performance of the SEM-Ctrl control group and the NonSEM-Trt treatment group were compared. This comparison allows consideration of which factor may have contributed more to the SEM-Trt comparative success over each of these groups. The SEM-Ctrl and NonSEM-Trt groups each had a disadvantage with respect to the SEM-Trt group:

1. Although the SEM-Ctrl students were also science and engineering majors (SEM) who had similar mathematics aptitudes (see Table 2) and mathematics coursework (e.g., precalculus and calculus courses), they were disadvantaged because during the semester they performed less demanding experiments, Level II, compared to the Level III of SEM-Trt students.

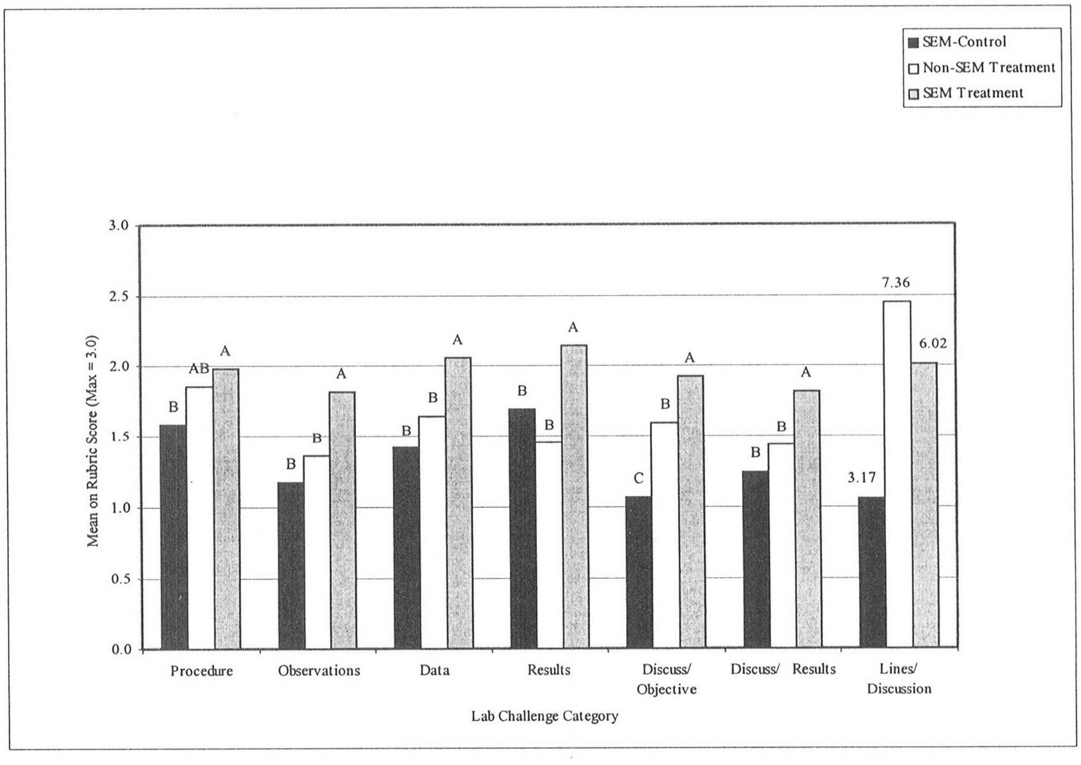
2. Although the NonSEM-Trt group was exposed to the same instructional method (Level III) as was the SEM-Trt group, these students majored in programs, such as nursing or applied sciences, with lower levels of required mathematics coursework (e.g., college algebra) and less mathematical emphasis in their lecture course.

The results (see Figure 1) showed that the NonSEM-Trt students, on average, were better able to discuss whether the objectives were accomplished ($p = 0.002$), compared to the SEM-Ctrl students. In addition, they ($M = 7.36$ lines, $SD = 2.55$) wrote a significantly longer discussion ($p < 0.001$) in terms of number of lines, when compared to SEM-Ctrl group performance ($M = 3.17$ lines, $SD = 2.72$). On the other hand, the SEM-Ctrl students tended to obtain and record slightly better results ($p = 0.07$) than did the NonSEM-Trt students. This tendency may have occurred because the SEM-Ctrl group logged almost twice as many hours of hands-on laboratory work (3 hours weekly over two semesters), compared to the NonSEM-Trt group (3 hours weekly over one semester).

Development of Laboratory Competence

To address Research Question 2, the relative proportions of students in the treatment group (SEM-Trt) and control group (SEM-Ctrl) who developed "laboratory competence" were determined. The results revealed that most (72.5%) SEM-Trt students developed laboratory competence at the mid-range standard (i.e., = 2 points on five of six categories), whereas only a few control-group students (30.5%)

Figure 1. T-*Test comparison of student responses to the Laboratory Challenge categories for the two treatment groups(SEM-Trt, Non-SEM-Trt) and the control group (SEM-Ctrl). Groups with the same label (A, B, or C) are not significantly different at* p < .05.



developed this level of competency. These results demonstrate that the innovative laboratory teaching approach used by the SEM-Trt group (Level III) was much more effective at developing a coherent set of laboratory investigative skills than was the SEM-Ctrl verification/cookbook approach (Level II). However, when all three groups were compared at a higher standard of competency (i.e., = 3 points on five of six categories), essentially no difference was found between the two SEM student groups. Specifically, the numbers of fully competent students were as follows: 4 SEM-Ctrl students (6.8%) and 5 SEM-Trt students (9.8%), but no NonSEM-Trt students met this standard. Also, although 1 NonSEM-Trt student just missed being classified as fully competent, it appears that nursing and applied science majors may lack the combination of interest in science, mathematics aptitude, and prior chemistry knowledge demanded by this higher standard of laboratory competency.

Qualitative Probe: Student Awareness of Thinking Skills

Qualitative research approaches can reveal a student's thinking process at a deeper level than can statistical methods over a large number of students. Regarding the essay question designated for Research Question 3, (see Qualitative Research Method), many students selected the laboratory challenge as the activity that encouraged them to think rather than to memorize and recite the answer. Apparently, most SEM-Trt students were not accustomed to a "thinking cycle" in their normal coursework, that is, a cycle in which they formulate an idea, test it, obtain feedback, and then modify the idea if the feedback extended or negated the idea. To quote one student:

> Well, of course the lab challenge made me use my brain, which I found was not [that] much use to me, but it was challenging trying to use my own memory and skills to figure out a problem.

Chemical educators have recognized the difference between a real problem versus an exercise (Bodner, 1987). With the former, students may read the problem statement but may not know initially what to do first, whereas an exercise triggers their algorithmic memory of a previously worked problem. Many SEM-Trt students described the laboratory challenge as being a "real problem," for example,

> Yes, the lab challenge induced me to use my thinking skills. At first, I had no idea of what to do, but after thinking for a while, I came up with different ideas. It was like a case that needed to be solved.

Another student expressed the view that the laboratory challenge induced her to relate the challenge task to other lab activities over the entire semester: "The lab challenge made me think about techniques and methods used in previous experiments (even experiments that I had done in Lab 101 previously!)." Thus, she even related her laboratory experience from a previous semester to the task of designing the challenge experiment. However, she was probably not used to thinking in this manner, because she wrote, "This [laboratory challenge] was frightening! I finally decided what procedure to use but felt very unsure of my method." A third student expressed the idea that scientific reasoning/thinking skills developed throughout the semester culminated in the laboratory challenge: "Yes. Most of the labs required some kind of thinking or planning. The lab challenge was the toughest. It was much more difficult to plan a lab from the very beginning than to plan with a predetermined outline."

Thus, as illustrated by these responses, at least some SEM-Trt students were aware of the demand placed on their thinking skills during their encounter with the Laboratory Challenge Experiment. This type of awareness has been shown to be a necessary mental process when people are successfully engaged in complex problem-solving situations.

Discussion and Implications

Discussion of Quantitative Results

Students were challenged during these laboratory practical examinations because they were faced with a "procedureless" laboratory activity, in which they had to synthesize knowledge and skills taken from several previous laboratory tasks. They had to write a laboratory report without outside assistance that communicated what they planned, what they did, what they saw, and what it meant. With respect to Research Question 1a, although SEM-Trt students struggled initially to understand the problem, this study's results suggest that they were better able to cope with a procedureless laboratory activity than were their SEM-Ctrl cohorts.

This finding suggests that students who are accustomed to performing verification experiments (SEM-Ctrl) are not able to communicate the results of their experiment (Hilosky et al., 1998). Surprisingly, the SEM-Ctrl group was not able to obtain and record good results from a chemical task (Figure 1), compared to the SEM-Trt group. One would expect the SEM-Ctrl students to excel on this skill because they had a double dose of hands-on laboratory experience (i.e., 28 laboratory periods), compared to the SEM-Trt students (14 laboratory periods). However, the opposite effect was found; the SEM-Trt group, on average, obtained and recorded better results than did their SEM-Ctrl cohorts ($p = 0.001$). The double dose of lab experience produced a tendency for SEM-Ctrl students ($p = 0.07$) to outperform NonSEM-Trt students on this skill. Overall, the SEM-Ctrl verification-based laboratory program was unable to transform "hands-on" skills into other investigative skills, despite a biased comparison made between two populations—one with more rigorous academic majors in SEM-Ctrl, compared to those in the NonSEM-Trt group.

For Research Question 2, level of laboratory competence was determined by re-analyzing the quantitative data (Figure 1) from a different perspective. The results indicated that most SEM-Trt students (72.5%) attained the midrange competent level, but only 30.5% of SEM-Ctrl and 28.6% of NonSEM-Trt students attained this same level. These results suggest that the greater extent of competency found among SEM-Trt group students may have resulted from the interaction of two factors. That is, they possessed an initial factor (Factor 1)—good scholarly characteristics (i.e., higher ACT-mathematics aptitude and better prior chemistry knowledge)—that were combined with a second factor (Factor 2)—learning by means of an instructional approach that encouraged them to employ cognitive skills that were explicitly taught during the semester (Goh et al., 1989). Consequently, the comparative lack of competency found in the other groups can be explained as follows: Fewer SEM-Ctrl students attained competency because they had not experienced Factor 2, while fewer NonSEM-Trt students possessed Factor 1. On the other hand, results at the fully competent level reinforced a belief that some chemistry teachers have expressed—the brightest students (top 10% in SEM-Ctrl or SEM-Trt) will find a way to become successful regardless of the instructional method they experience.

Discussion of Qualitative Results

For Research Question 3, SEM-Trt written responses to an essay question on the final laboratory examination were analyzed to determine whether students were aware of their investigative skill development. This question probed their awareness of the thinking skills they used during the laboratory challenge experiment. SEM-Ctrl students were not asked these essay questions, as was pointed out earlier, due to the rigid multiple-choice format of their written final laboratory examination.

Some SEM-Trt students indicated that the Laboratory Challenge Experiment posed a real problem or puzzle to be solved rather than just an exercise, which reflects the real problem-exercise distinction made by several researchers (Bodner, 1987; Lythcott, 1990). It was challenging because they initially "had no idea of what to do," and because, as one student stated, he had to "use my own memory and skills to figure out a problem." Each student had to generate or modify a procedure that would produce data that might solve the problem.

This generative process is consistent with constructivist theory (Shiland, 1999), and it is illustrated by a student's remark regarding the Laboratory Challenge Experiment, which "made me think about techniques and methods used in previous experiments." A synthesis of these individual statements suggests that these students were saying, in effect, that the real challenge was devising, executing, and then explaining in writing a procedure for conducting a scientific investigation. This process combines a chemical task with a writing task so that students use their own creative ideas, and thus, they should begin to appreciate the investigative skills they have developed (Goodman & Bean, 1983).

Implications for Science Laboratory Instructors

Over several decades, many high school and college administrators have questioned the value of chemistry laboratory programs primarily due to economic factors (Lagowski, 1989). That is, these programs incur high costs for the purchase of chemical supplies and equipment, increasing costs for the disposal of these materials, and potential liability risks due to student use of hazardous chemical substances. Instructors need convincing classroom-based data to counter these concerns. A laboratory practical examination, as described in this study, can serve as a measuring instrument that documents the value or impact of a laboratory program in helping chemistry students develop their investigative skills. Although some of these skills are specific to chemistry (e.g., manipulative skills), others are generic skills, (e.g., communication skills and observational or inferential skills), which are transferable to other academic programs.

Another implication for laboratory instruction from these results is that the required competence level of an instructional approach (Levels I to IV) should initially match the scholarly characteristics of students. Two types of mismatches are possible: the instructional demand may be greater than students' capabilities and vice versa. For example, if instructors use an open inquiry method (Level IV+) with first-year high school chemistry students, then students may become easily overwhelmed and thus experience frustration rather than enlightenment (Ealy & Ealy, 1994). Conversely, when instructors use the verification approach (Level I or II) in a college laboratory course designed for science and engineering majors (SEM), then their students may be underchallenged and, thus, unable to develop and demonstrate investigative skills, as demonstrated by the results of the SEM-Ctrl group in this study. This problem may be pervasive because most general-chemistry laboratory programs in the United States are taught using a verification approach (Abraham et al., 1997; Hilosky et al., 1998*)*. If a program elects an inquiry-based approach, then instructors must work hard to develop and implement such a radical curricular change (Huber & Moore, 2001). Fortunately, another strategy is to make modifications in some verification-type activities and to transform them into investigative experiments (Herman, 1998).

References

Abraham, M. R., Cracolice, M. S., Graves, A. P., Aldhamash, A. H., Kihega, J. G., Palma-Gil, J. G., & Varghese, V. (1997). The nature and state of general chemistry laboratory courses offered by colleges and universities in the United States. *Journal of Chemical Education, 74,* 591-594.

Beasley, W. (1991). Matching laboratory learning goals to evaluation of student performance: A standards-based approach. *Journal of Chemical Education, 68,* 590-591.

Bodner, G. M. (1987). The role of algorithms in teaching problem solving. *Journal of Chemical Education, 64,* 513-514.

Bybee, R. W., Ferrini-Mundy, J., & Loucks-Horsley, S. (1997). National standards and school science and mathematics. *School Science and Mathematics, 97,* 325-334.

DeMeo, S. (1997). Does copper metal react with acetic acid? *Journal of Chemical Education, 74,* 844-846.

DiBiase, W. J., & Wagner, E. P. (2002). Aligning general chemistry laboratory with lecture at a large university. *School Science and Mathematics, 102*(4), 158-177.

Domin, D. S. (1999). A review of laboratory instruction styles. *Journal of Chemical Education, 76,* 543-547.

Ealy, J. & Ealy, J. (1994). Frustration + cleverness ≠ learning. *Journal of Chemical Education, 71,* 148-149.

Goh, N. K., Toh, K. A., & Chia, L. S. (1989). Use of modified laboratory instruction for improving science process skills acquisition. *Journal of Chemical Education, 66,* 430-432.

Goodman, W. D., & Bean, J. C. (1983). A chemistry laboratory project to develop thinking and writing skills. *Journal of Chemical Education. 60,* 483-485.

Herman, C. (1998). Inserting an investigative dimension into introductory laboratory courses. *Journal of Chemical Education, 75,* 70-72.

Hilosky, A., Sutman, F., & Schmuckler, J. (1998). Is laboratory based instruction in beginning college-level chemistry worth the effort and expense? *Journal of Chemical Education, 75,* 100-104.

Horton, R. E. (1928). Does laboratory work belong? *Journal of Chemical Education, 5,* 1432-1433.

Huber, R. A., & Moore, C. J. (2001). A model for extending hands-on science to be inquiry based. *School Science and Mathematics, 101*(1), 32-42.

Keys, C. W. (2000). Investigating the thinking processes of eighth grade writers during the composition of a scientific report. *Journal of Research in Science Teaching, 37,* 676-690.

Lagowski, J. J. (1989). Reformatting the laboratory. *Journal of Chemical Education, 66,* 12-14.

Lagowski, J. J. (1990). Entry level science courses: The weak link. *Journal of Chemical Education, 67,* 185.

Laws, P. M. (1996). Investigative work in the Science National Curriculum. *School Science Review, 77*(281), 17-25.

Lock, R. (1990). Open-ended, problem-solving investigations: What do we mean and how can we use them? *School Science Review, 71*(256), 63-72.

Lythcott, J. (1990). Problem solving and requisite knowledge of chemistry. *Journal of Chemical Education, 67,* 248-252.

Meester, M. A. M., & Maskill, R. (1995). First-year chemistry practicals at universities in England and Wales: Aims and the scientific level of the experiments. *International Journal of Science Education, 17,* 575-588.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academy Press.

Ryder, J., & Leach, J. (1999). University science students' experiences of investigative project work and their images of science. *International Journal of Science Education, 21*, 945-956.

Shiland, T. W. (1999). Constructivism: The implications for laboratory work. *Journal of Chemical Education, 76,* 107-109.

Silberman, R., Day, S., Jeffers, P., Klanderman, K., Phillips, M. G., & Zipp, A. (1987). Unusual laboratory practical examinations for general chemistry. *Journal of Chemical Education, 64,* 622.

Editors' Note: The author would like to acknowledge the support of the Department of Chemistry at McNeese State University with special thanks to Don Coombe who worked with the author to establish the "new" laboratory program. The assistance of Robert DeMay in the data analysis was invaluable. Also, the comments and suggestions of David Pringle, Henry Heikkinen, and three anonymous reviewers greatly improved this article.

Correspondence concerning this article should be addressed to Jerry P. Suits, Department of Chemistry and Biochemistry, University of Northern Colorado, Greeley, CO 80639.

Electronic mail may be sent via Internet to jerry.suits@unco.edu

Appendix

Scoring Rubric to Assess Performance on Six Investigative Skills of the Laboratory Challenge Experiment.

Procedure

3 Someone else could follow the procedure and do the entire experiment from it
2 Some steps are clear but others are not
1 Vague
0 None

Observations

3 Inferences and observations linked together
2 Inferences and observations discussed separately
1 Observations- no inferences included
0 None

Data

3 Very specific
2 Generic
1 Vague
0 None

Results

3 Specific results- accurate
2 Generic results
1 Vague
0 None

Discuss Objective (accomplished?)

3 Objective and specific support from procedure and data collected
2 Objective and generic information
1 Vague
0 None (section left blank)

Discuss Results (observed/expected)

3 Theoretical and observational linked together
2 Theoretical- what it should be
1 Observational-what it was
0 None