

RUNNING HEAD: FEEDBACK IN CAI

Computer Assisted Instruction: The Role of Feedback in Performance and Academic Self-  
Efficacy

Diann E. Mazingo

University of Colorado at Denver

EDLI7712

Learning Theories, Dr. W. Goodwin

REM7120

Advanced Methods of Quantitative Inquiry and Measurement, Dr. L. Goodwin

Spring 2003

### Abstract

Feedback has been identified as a key variable in developing academic self-efficacy (ASE). The types of feedback can vary from a traditional, objectivist approach that focuses on minimizing learner errors to a more constructivist approach, focusing on facilitating understanding. The influx of computer-based courses, whether online or through a series of computer-assisted instructional (CAI) modules, requires that the current research of effective feedback techniques in the classroom be extended to these computer environments in order to impact the instructional design of these learning experiences. In this study, gender and exposure to different types of feedback during a chemistry CAI module were studied in relation to ASE and performance on an objective-driven assessment (ODA) of the chemistry concepts covered in the module. No significant changes in ASE across time were found. Also, no significance of the between-subjects or within-subjects effects for the ODA was observed. These findings are discussed in relation to the need to further refine the instruments to continue exploring the possibility that the type of feedback might play a role in developing ASE, and consequently, academic performance. Future research, building on this pilot, may reveal significance that could have impacts on the instructional design of the growing body of online and computer-based coursework.

Computer Assisted Instruction: The Role of Feedback in  
Performance and Academic Self-Efficacy

*Background*

*Academic feedback.* Academic feedback, the procedures a teacher uses to provide students with information on the accuracy of their oral or written responses to academic questions, has been found to be “strongly and consistently related to student achievement” (Filby & Cahen, 1985). There is a great deal of research on the use of academic feedback in the classroom, focused primarily on creating effective feedback opportunities, question types and delivery, and instructor’s reactions to student responses. However, the advent of the technology age, bringing online education and computer assisted instruction (CAI) tools, requires new investigations of how academic feedback is most effective in technology based instruction environments. These computer-based environments offer new venues for both immediate and delayed feedback, and new research is needed to assess the efficacy of different types of feedback as they apply in these new situations.

Research on feedback in the classroom environment explores many different aspects of generating effective interactions with students. Classroom environments and differences in the types and delivery of feedback from the instructor both contribute to the likelihood of successful feedback interactions that improve student achievement. Teacher efficacy, judged by the performance of a teacher’s students, has been correlated with the quality and quantity of questions asked. More effective teachers provide several times more opportunities for students to answer questions than less effective teachers (Brophy & Good, 1986).

In order to create environments where academic feedback can be delivered successfully, teachers must consider their classroom time-management and expectations and modify them

appropriately to facilitate increased opportunities to provide adequate feedback, both immediate and delayed. Restructuring the classroom environment to free up time for instruction by minimizing distractions from behavior and misconduct allows for more feedback opportunities. Also, reorganizing the curriculum to present it in smaller segments enables more opportunities to check for understanding. Furthermore, the variety of questions that an instructor poses to students can affect student achievement as well. Effective teachers carefully consider such factors as question difficulty and clarity when providing feedback to students. Contrary to the common notion that higher-level questions, often involving inductive reasoning, are better than lower-level questions, research has shown that high success rates for answering questions of any level positively affect student achievement. However, if the questions posed are not sufficiently clear, success will decrease. Factors that reduce the clarity of questions may include (a) vague or ambiguous questions, (b) disjointed questions, (c) multiple questions at once, and (d) failing to get the students' attention before posing the question. The delivery of questions, therefore, involves a complex monitoring of the environment, content, ability levels, and objectives (Brophy & Good, 1986).

*Online Academic Feedback.* The development of online courses changes common perceptions of feedback in terms of the types of feedback available and opportunities for providing feedback. Instructional designers of CAI modules often take for granted the process of providing feedback by assuming that feedback should always be provided promptly and unambiguously. However, if a more broad definition of feedback is adopted, such as “any action or presentation that is based on prior performance of the learner and whose purpose is to facilitate learning,” (Wilson, Cole, Myers, & Clariana, 1994) then it becomes necessary to acknowledge that there are other, less objectivist and more constructivist ways to provide

feedback in an online or CAI environment. The broadened definition of feedback opens up more possibilities to better simulate real-life interactions within the virtual world of a computer based learning experience. While there are many key feedback manipulables that an instructional designer can control, this study is interested in comparing the effects of objectivist vs. constructivist approaches to providing feedback during knowledge mastery performance items.

Objectivist feedback differs from constructivist feedback primarily in their contrasting goals of instruction. The primary purpose of objectivist feedback is to minimize learner errors. This more traditional feedback is often given in the form of praise for correct answers or statements such as “incorrect” or “your answer is wrong” for mistakes made by the learner. Brooks and Brooks (1993) stated that these blanket rejections of student responses lead to a lessened desire to think about and explore issues. Instead, if a learner encounters a more constructivist feedback approach, where the primary purpose is to facilitate understanding, s/he is more likely to engage in thinking processes that construct new understandings and acquisition of new skills.

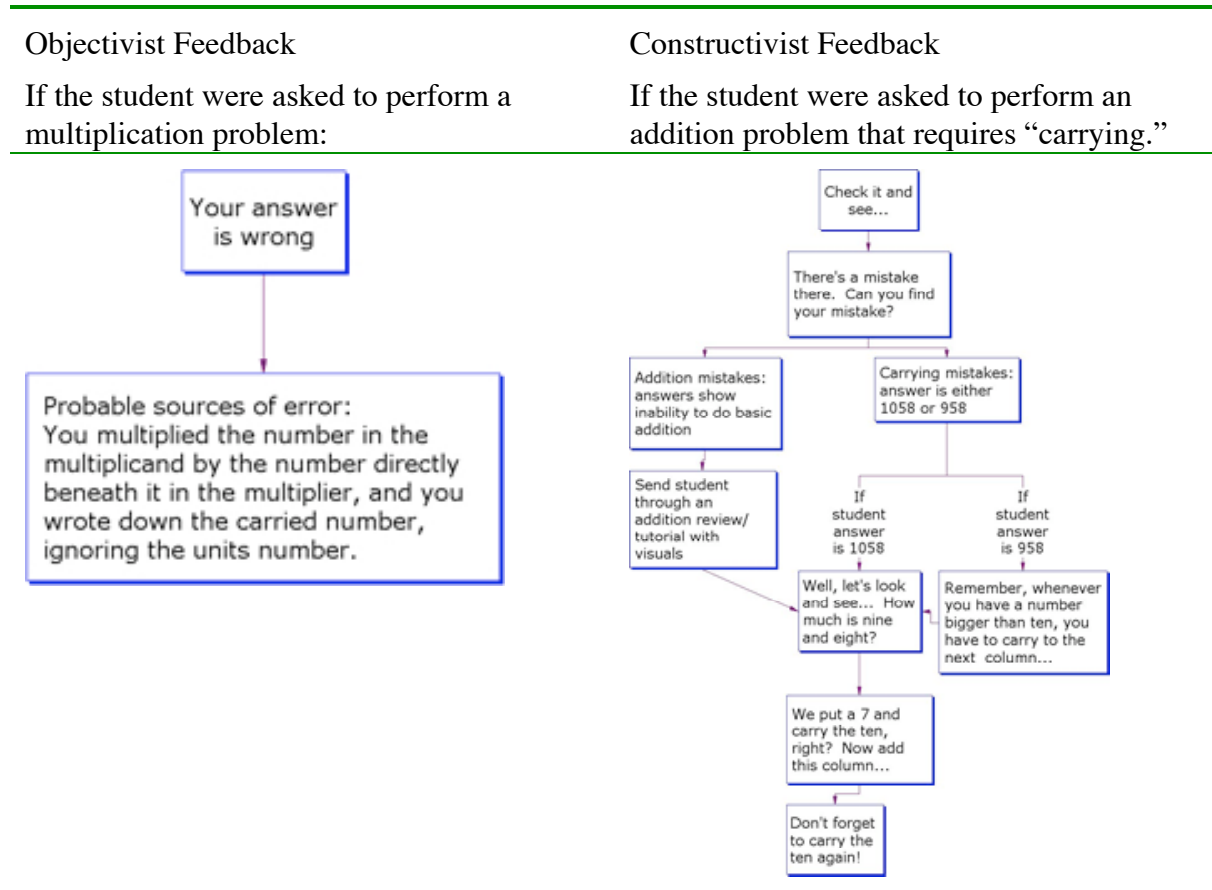
Both immediate and delayed feedback are essential for fostering adult learning and comprehension (Gagne, 1984). Immediate feedback refers most often to verbal feedback given at the time of the question and answer. Delayed feedback is more commonly found as written feedback to a written assignment or assessment. Feedback in the online environment either can be immediate or delayed, depending on the structure and design of the course or assessment. For instance, online tests can be given so that each question is graded with feedback provided prior to advancing to the next question. Another possibility is that the assessment is graded after completing all questions. A third possibility involves a delay between submitting the assessment electronically and receiving the grade as feedback.

Computer-based courses or modules also open up possibilities for feedback that may not exist in more traditional classrooms by enabling communication between students and the instructor, whether the instructor is the actual person behind the course or the feedback programmed into the learning activities of the course. Individual feedback is more powerful than group feedback (Archer-Kath, Johnson, & Johnson, 1994). This finding is not surprising because it is in line with other academic constructs, such as self-efficacy, in that individualized attention increases student performance. Especially at the collegiate level, instructors and students are not given the environment to interact directly with each other because of class sizes, structures, and preexisting traditions of lecture-based instruction. In an online environment, students have more direct access to an instructor to ask individual questions and possibly receive individual feedback, and therefore possibly have a greater impact on the academic self-efficacy of the learners.

However, there are more feedback opportunities in computer-based courses than just during assessments and forms of asynchronous communication, like email and discussion boards. Similar to face-to-face classroom experiences, effective instructors must modify the environment and content to provide opportunities for feedback to occur. In this manner, feedback in an online or CAI setting can be distributed into learning activities and not just during formal assessments and written communication. Pioneer work in differing computer-based learning environments by Lepper and Chabay (1988) explored the process of using differentiated feedback depending on the learner response for diagnosing learner errors and providing appropriate constructivist-like feedback to help guide the learner through the cognitive processes necessary to solve the original problem. This method was innovative in computer-based settings because previous feedback was often limited to an objectivist-type of feedback utilizing

statements of “correct” or “incorrect,” with incorrect responses sometimes followed by a list of possible errors for the learner to revisit.

**Figure 1:** Lepper and Chabay (1988) feedback differences for CAI



Their model, which was intended to produce feedback that maximally motivates, in addition to delivering corrective information, resembles constructivist instruction in its sequence and delivery because students come to their own understanding with the strategic feedback provided by the computer program (Figure 1). Therefore, the learner immediately receives information as to her or his success at answering the question; and, if the answer was incorrect, a series of prompts lead the learner to develop an understanding of her or his mistake(s).

*Connections to Academic Self-Efficacy*

*Self-efficacy.* Providing feedback is connected to the construct of self-efficacy. Self-

efficacy is the set of beliefs about one's capabilities to learn or perform at designated levels (Bandura, 1994). These beliefs directly affect a person's ability to persevere and ultimately succeed at a given task. There are many factors that influence the development of self-efficacy. First of all, personal mastery experiences positively impact an individual's self-efficacy because previous success at a given task raises the individual's perception of her or his ability to accomplish the task again. Even if the two tasks are not directly related, it is possible for success at something the individual determined was difficult would encourage the individual to tackle other perceived difficult tasks. Second, vicarious experiences play a role in the development of self-efficacy. If someone for whom an individual identifies as being similar to him is successful at a given task, he is more likely to determine that he or she has a likelihood of success as well. Third, social persuasion in the form of verbal or written communication increases an individual's self-efficacy, especially if the persuasion is realistic to the individual's abilities and/or talents. Finally, somatic and emotional states, or how emotional and physical reactions to certain activities are interpreted, can positively or negatively influence self-efficacy perception. Since CAI modules are generally completed in isolation from other learners, most of the self-efficacy changes in an individual will result from mastery experiences and social persuasion. Including feedback within a mastery task that encourages learners to develop their own understanding encourages both of these factors.

*Impact of self-efficacy.* Levels of self-efficacy contribute to a person's choices, effort, persistence, resilience, and achievement (Bandura, 1997). Numerous examples show how people, in the face of rejection, continue trying and eventually succeed. For example, Thomas Edison failed one thousand times before successfully inventing the light bulb. Another example of perseverance are Tom Landry, Chuck Noll, Bill Walsh, and Jimmy Johnson accounted for 11

of the 19 Super Bowl victories from 1974 to 1993. They also share the distinction of having the worst records of first-season head coaches in NFL history in that they did not win a single game (Pajares, 2001). These are just a couple of testaments to the idea that people with high self-efficacy will choose to continue exerting effort towards a particular achievement and ultimately, succeed because of their persistence and resilience.

While self-efficacy has been studied in detail for a wide range of behaviors and situations, it is especially important when exploring academic success. The development of academic self-efficacy (ASE) is complex in the sense that many different people and situations influence its development. Familial, peer, social, and schooling influences play overlapping and fundamental roles in the initial development of ASE, however, this construct changes over time and once a child matures into adulthood, many opportunities still exist for continued development and enhancement of ASE. Unfortunately, colleges and universities are often structured so that students miss out on important opportunities for continuing to foster their ASE. While mastery experiences are the primary influence on self-efficacy, and post-secondary teachers generally take responsibility for providing intellectual challenges, developing confidence is also the result of the actions of others. Emotional support and encouragement from others is a very powerful way to provide increases in self-efficacy. However, many colleges and universities do not emphasize the instructor's role in nurturing students' self-beliefs. In a recent survey of college seniors at Emory University, 97 percent said that Emory professors provided them with intellectual challenges and stimulations. Only 20 percent said they received frequent emotional support and encouragement (Pajares, 2000). Thus, students received support in developing ASE through mastery experiences, but lacked support for the increases provided by the social persuasion of others.

*Academic self-efficacy.* The construct of self-efficacy is used in a variety of ways related to the academic setting. ASE is often investigated when researchers look at differences in academic performance across varying content areas and different ethnic, gender, and social groups. Researchers often investigate the different factors that affect the development of ASE. Finally, ASE is most commonly used as a measure to predict academic achievement.

Self-efficacy is known as a context-specific construct (Zimmerman, 1995). Numerous studies exist that have investigated the differences in ASE within specific content areas (Bong, 2002; Joo, Bong, & Choi, 2000; Marsh, 1992; Marsh, Walker, & Debus, 1991; Pajares & Miller, 1994, 1995, 1997). These studies suggest that not only do the predictive abilities of ASE measures increase when the measure focuses on one content area and performance within that area; they also imply that the self-efficacy-outcome links are stronger within the same domain than across different domains (Joo et al., 2000). Bong (2002) took this facet of ASE measurement a step further and investigated three different levels of specificity in two different subjects to analyze any cross-domain interactions. This study allowed Bong to add further support to the argument for specificity within a context domain, and it allowed her to test whether self-efficacy actually demonstrates stronger relationships with performance measures in the same subject area than with performance in a different area. Bong concluded that self-efficacy perceptions in a specific school subject were content specific to achievement. In other words, English self-efficacy predicted only English achievements and math self-efficacy predicted only math achievements. Cross-domain predictions were weak and not statistically significant.

Gender and ethnic differences in measures of self-efficacy is another facet of ASE that is commonly examined. Investigations of ASE gender differences are extensive. Initial research

into gender differences between math, science, and language arts concluded that boys and men tend to be more confident than girls and women in math and science while girls and women are more confident in the language arts than boys and men (Meece, 1991; Pajares & Miller, 1994; Wigfield, Eccles, & Pintrich, 1996). However, further research has revealed that these conclusions are influenced by myriad factors, including previous achievement, age/development, and preexisting stereotypes. The differences between male and female ASE is greatly reduced when researchers control for previous achievement (Pajares, 1996). However, because previous achievement often depends on ASE, this argument is circular. The age and development of the respondents also influences the findings of ASE measures. Little evidence suggests ASE differences between genders of elementary-aged children. The reported differences do not emerge until adolescence (Wigfield et al., 1996) when girls show a decline in self-efficacy. Finally, gender reporting of ASE is influenced by preexisting stereotypes of the relative capabilities of males and females. Researchers have found that differences in course and career selection and in confidence due in part to the role of stereotypes (Pajares & Valiante, 1999).

Research on ethnic and race differences in perceived ASE is not as extensive as research on gender differences. However, various studies have shown that minority students have lower ASE than nonminority students but that these differences do not remain when socioeconomic status is controlled (Graham, 1994). Additionally, positive relationships have been found between Black consciousness, an individual's beliefs or attitudes about her or his self, own race, and the White majority vis-à-vis the Black experience, and ASE (Okech & Harrington, 2002). Further research into ethnic differences is suggested to understand the differences between ASE in a broader sampling of the different ethnic and racial backgrounds of students at the elementary, secondary, and collegiate levels.

Another area of focus for ASE research investigates the different factors that affect levels of ASE. Most of this research appears to be at the post-secondary level and often focuses on how different instructional strategies impact self-efficacy. A study at Indiana-Purdue University at Fort Wayne investigated the effects of a communication designed specifically to enhance the self-efficacy of introductory psychology students (Jackson, 2002). Through the use of email, the instructor provided efficacy-enhancing or neutral replies and monitored the effect of these communications on test performance. Self-efficacy beliefs were both significantly related to exam scores and significantly affected by the efficacy-enhancing communication. Another study at the collegiate level analyzed the effects of reciprocal peer tutoring (RPT) on self-efficacy and exam performance (Griffin & Griffin, 1997). While previous research indicated that RPT positively influences achievement and reduced participants' level of stress and anxiety, the Griffin and Griffin study showed no significant differences between RPT and non-RPT group performances on academic measures or ASE. Finally, a growing body of research focuses on developing computer self-efficacy (Albion, 2001; Joo et al., 2000; Madorin & Iwasiw, 1999; Russell, 1999; Smith, 2001). These studies all point to mastery experience and time using computers as the main sources for developing computer self-efficacy. Since computer-use is an integral part of today's educational platforms, continuing research on developing computer self-efficacy in both learners and instructors is necessary.

Finally, ASE has direct influence on levels of academic achievement. "Experimental and correlational research in schools suggests that self-efficacy is positively related to a host of positive outcomes of schooling such as choice, persistence, cognitive engagement, use of self-regulatory strategies, and actual achievement" (Linnenbrink & Pintrich, 2002). Numerous studies investigate the correlation between ASE and exam performance. House (2000a, 2000b)

focused his research on how academic background and self-beliefs can serve as predictors for performance in science, engineering, mathematics, and health science majors. He found that self-beliefs accounted for 20% of the variance in students' cumulative grade point averages. Research on self-regulated learning is also closely tied to ASE and suggests that students with high efficacy are more apt to be successful in self-regulated learning environments (Miller, 2000; Pajares, 2002; Zimmerman, 2002). This area of research also connects to differences in gender and ASE because, in general, girls have more goal-setting and planning strategies, keep records, and self-monitor more frequently than boys, lending them a higher self-efficacy for those tasks (Pajares, 2002). Research on the malleability of self-efficacy beliefs and grade goals as predictors of exam performance (Vrugt, Langereis, & Hoogstraten, 1997; Wood & Locke, 1987) continue to confirm other bodies of research that positively correlate levels of self-efficacy to levels of achievement.

### *Purpose*

The purpose of this study is to investigate the effects of receiving constructivist vs. objectivist type feedback during a CAI module on levels of ASE and performance on an objective-driven assessment (ODA). Both gender and treatment are between-subjects factors while "time" (pretest vs. posttest) is a within-subjects factor. There are two dependent variables, ASE and ODA. Based on previous research of both classroom and computer-based environments, feedback, and ASE my hypotheses are that learners receiving the constructivist feedback approach will show significant mean gains in both dependent variables than the learners receiving objectivist-type feedback.

### *Method*

*Subjects.* Subjects consisted of 77 honors chemistry students at Eaglecrest High School in Aurora, Colorado. The students participated in the study as a required assignment for their class with a grade based on completion of all parts of the study. Participants were stratified by gender and then randomly assigned to one of two treatment groups. The final number of subjects for which data are available only numbered 44. The loss of subjects was the result of absences from one or more parts of the study or the unpredictable and uncontrollable loss of pre-post data stored on an outside server. In the two groups of subjects, 50% of the students were male and 50% were female. Ages ranged from 15-18 and all were high school sophomores, juniors, and seniors. The control group ethnicity breakdown was 73% white, 14% Asian, 9% Hispanic, and 4% African American. The experimental group was 91% white and 9% Asian. Because of the random assignment, the pronounced difference in the ethnic breakdown of the groups was surprising. However, since ethnicity was not a factor of this design, and these values are provided as qualitative information only to help understand the limitations of the study for generalizing to a broader population.

*Design.* Two separate repeated-measures factorial designs were used to analyze the data. Gender and group were the between subjects factors and the pretest/posttest data composed the within subject factor, labeled “time.” The gender-stratified, random assignment to one of two treatment groups with a pretest/posttest repeated measures design is graphically depicted to show the separate two-way repeated measures analysis of variance (RM-ANOVA) tests for (a) average ASE and (b) percentage score on the ODA (Table 1). This method was chosen over the less powerful two-way ANCOVA using the pretest as a covariate because an analysis of the pretest and posttest means verified that the random assignment was effective. Thus, using the pretest as

a covariate was unnecessary to eliminate selection bias and would, therefore, only serve to increase power.

**Table 1:** Graphic depiction of the experimental design for both dependent variables

Gender	Treatment	<i>Score on Dependent Variable</i>		
		Ss	Occasion (Time)	
			Pretest	Posttest
Female	Control	1		
		2		
		3...		
		11		
	Experimental	1		
		2		
		3...		
		11		
Male	Control	1		
		2		
		3...		
		11		
	Experimental	1		
		2		
		3...		
		11		

There are several advantages to repeated measures designs. First of all, subjects serve as their own control, comparing subjects to themselves, which leads to greater efficiency of economy (power) because the between-group variance is usually bigger than within-group variance. Second, the design controls for subject heterogeneity, or the innate differences between people. Disadvantages of repeated measures designs include (a) carry-over effect, if one treatment is administered before the previous treatment has completely worn off, (b) latent effect, where treatment activates the dormant effect of another, and (c) order/learning effect, similar to the testing effect where the response on dependent variable might be different because the subjects have taken the measure before.

In this particular study, the advantages are great strengths since each individual student enters the study with different innate abilities and levels of previous chemistry performance successes. Therefore, since each subject is compared only to herself or himself, the generalizability of the results is greater than if the subjects were part of a larger group mean comparison. The typical disadvantages to this analysis are not large concerns for this study. Since there is only one treatment, the study will not be hindered by carry-over effects. Additionally, the latent effect is not applicable since it is unlikely that there were any dormant effects present. The only disadvantage that is directly related to this particular study is the order/learning effect. Since the same exact questions were used for both pretest/posttest pairs, it is possible that an individual would answer the posttest differently having been previously exposed to the same questions.

The assumptions for RM-ANOVAs are similar to those for a two-factor mixed-effects model and include (a) homogeneity of variance, (b) normality of population, (c) compound symmetry/sphericity, and (d) equality of covariance matrices. Violations of the homogeneity of variance would most likely result in an increased likelihood of a Type I and/or Type II error. This assumption was not violated for these data because there are equal  $n$ 's in all cells. Likewise, normality of the population is not a violation concern because when there are equal  $n$ 's, violation has minimal effects. Neither is sphericity a concern since there are only two between-subject factors, the pretest and the posttest. Therefore, there is only one possible pairwise correlation making sphericity a non-issue. However, the equality of covariance matrices, which states that the covariances between the scores of the subjects across the levels of the repeated factor ("time") are constant, is a concern. Since this analysis of variance is not robust to violation of this assumption, it is often violated. If this assumption is violated,

evidenced by Box's test of equality of covariance matrices, then there are three options available: (a) limit the levels of the within subjects factor to two, (b) use the adjusted F-tests, and (c) use a MANOVA analysis which has no compound symmetry assumption. Since my design only includes two levels of "time," violation of this assumption is not a concern.

*Instruments.* The author wrote the Colligative Properties of Solutions computer assisted instruction (CAI) module using Macromedia Flash. The Flash program facilitated learner interaction that controlled the computer-generated text, graphics, concept check questions, and feedback that focused on imparting mastery of four learner objectives about colligative properties. These four objectives were to (a) explain on a particle basis why a solution has a lower vapor pressure than the pure solvent of that solution, (b) explain on a particle basis why a solution has an elevated boiling point and a depressed freezing point compared with the pure solvent, (c) calculate the molality and mole fraction of a solution, and (d) calculate the boiling-point elevation of freezing-point depression of a given solution.

The module's four sections were similar in structure. Each section begins by introducing the learning objective, providing interactive examples demonstrating the concept(s) of the objective, and if there are mathematical calculations involved, animated step-by-step sample calculations. Immediately following this delivery of the concepts involved, learners proceed to an interactive "check your understanding" portion where students are given a sample problem and 4-5 multiple choice possibilities for the answer. Up until this point in the CAI module, the control and experimental groups receives identical treatment. The difference between the two groups is in the type of feedback offered upon selecting a correct or incorrect response in the "check your understanding" questions. Learners in the control group received a traditional objectivist-like feedback structure while the experimental group learners received more

constructivist-like feedback (Wilson et al., 1994). For example, selected screen shots in Appendix A illustrate the difference between the two feedback possibilities for one of the module's "check your understanding questions." Notice how the control group is given a blanket response of "incorrect, possible sources of error include..." while the experimental group is asked a series of leading questions that are intended to lead the learner through discovering her or his own understanding of the concept.

Two experts with an average of 12 years of advanced chemistry teaching experience were asked to review the technical content of the program for accuracy. Their suggestions helped correct a few minor flaws in calculations and clarify some of the explanations used in the module. One expert in the field of design theory also reviewed the program and aided in simplifying the navigation and making the module more intuitive to use.

Learners' content mastery of the four stated objectives was measured by a pre-post ODA multiple-choice test. A pre-post ASE instrument was used to measure learners' cognitive perceptions of behavioral performance. Finally, a qualitative self-response instrument measured satisfaction with the program.

Twenty-three knowledge test questions were compiled from the textbook resource materials. Each objective had a minimum of 4 specific questions that addressed the content from the module, with more difficult objectives having a greater number of questions on the instrument. An item analysis, based on the protocol from Hopkins (1998), of the posttest ODA data from both groups determined that certain questions yielded low item discrimination values. Of the 23 questions, 4 questions were discarded. Of the remaining 19 questions, 4 had *D*-index values indicating fair discrimination, 4 had good discrimination, and the remaining 11 items, with *D*-index values over 0.40, had excellent discrimination. Since there is a direct relationship

between item discrimination values and a test’s internal consistency reliability, discarding the 4 questions with low discrimination increases this instrument’s reliability. After discarding the 4 questions, a corrected split-half reliability coefficient of 0.84 was calculated, indicating that the instrument was highly reliable.

Examining a table of specifications for the proper balance of content and process objectives assessed the content validity of the instrument. Table 2 separates the 19 remaining ODA questions into their respective objectives and taxonomy level. The breakdown makes it obvious that the test represents the content and process objectives in proportion to their importance, a property of a test that is important for content validity. Therefore, the instrument used to measure the learners’ mastery of content is both reliable and valid.

**Table 2:** A Table of Specifications for ODA (Numerals in Parenthesis Refer to Specific Items on the Test)

Content Strata (Objectives/Topics)	Taxonomy Level		
	Knowledge	Higher	Total
1. explain vapor pressure	1 (1)	2 (8, 9)	3 (16%)
2. explain boiling-point and freezing-point	3 (2, 3, 13)	3 (6, 15, 16)	6 (32%)
3. calculate molality and mole fraction	3 (7, 18, 20)	2 (17, 19)	5 (26%)
3. calculate boiling-point elevation and freezing-point depression	2 (4, 5)	3 (21, 22, 23)	5 (26%)
Totals	9	10	19

The ASE measure used to investigate learners’ cognitive perceptions of behavioral performance also consisted of 23 items. Self-efficacy pertains to whether an individual attempts a task, persists, and achieves successful completion (Bandura, 1994; Bandura, Barbaranelli, Caprara, & Pastorelli, 2001; Cassidy, 2000; Pajares, 1996). Previous research has found poor correlations between increased levels of self-efficacy and knowledge, suggesting that each

construct measures a different component of learning (Bandura, 1994, 1997; Pajares & Miller, 1995; Schunk & Pajares, 2001; Zimmerman, 1995). Using Bandura's Guide for Constructing Self-Efficacy Scales (Bandura, 2001), 5 questions were developed from the program objectives, with each chemistry objective translated into a task-specific self-efficacy item related to the concepts required to master the objective. Additionally, 18 questions from Bandura's Children's Self-Efficacy Scale were selected from 3 main categories, (a) enlisting social resources, (b) academic achievement, and (c) self-regulated learning (Bandura, 2001). Learners were asked to indicate on a self-report response format the amount of confidence they had in their ability to perform each of the behaviors on a scale of 0-100. A 0 indicated that the learner thought they could not do the behavior at all and a 100 indicated that the learner was certain s/he could perform the behavior. A practice rating was administered beforehand to familiarize the learner with how to use the scale. The unconventional use of a 0-100 scale instead of a more traditional Likert-type scale was chosen based on research indicating that the predictive abilities of self-efficacy measures are increased with greater discrimination (Pajares, Hartley, & Valiante, 2001).

Construct validity was investigated using a factor analysis method. Using varimax rotation, six factors emerged from analyzing the posttest ASE data for all subjects. Further analysis of the factor loadings and the corresponding questions reveals groupings of questions that can be defined as (a) chemistry, (b) self-regulated studying, (c) general math, (d) written-word efficacy, (e) self-regulated planning and organizing, and (f) interactions with teachers and peers. The factor loadings were quite high, in general, with 18 of the 23 items loading at 0.700 or higher. While this analysis did not separate into the four groupings expected from the design, the groupings are still very justifiable upon further consideration of the questions. Furthermore, the posttest data for both groups were split into two parts yielding a corrected split-half reliability

coefficient of 0.91. This exceptionally high value indicates that the instrument has excellent consistency internal reliability.

The qualitative self-report instrument consisted of 6 questions addressing the learners' overall satisfaction with the CAI module experience. At this time, these data, in the form of free-response journal entries, are only used to help understand the quantitative findings and no formal qualitative analysis technique was used to interpret the results.

*Procedure.* The entire study took place over three days in a computer lab at the school. Each student was assigned to his or her own computer and the class was instructed to treat their environment as if there were no other students present. They were asked, therefore, to not speak with each other about the experience and could not ask questions of each other or the instructor. Each student was provided with specific, step-by-step, directions detailing how to login to the program and what order to complete the parts of the module. Since their participation grade for the study required that they adhere to these rules, the students complied with these restrictions.

After the initial directions were given, students were then released to complete the ASE and ODA pretests. Immediately following the completion of the pretests, students proceeded into the CAI module where they were encouraged, but not required, to take notes and be active learners throughout the objective instruction and concept checks. After completing the module, students finished the experience by completing both posttests and answering the 6 qualitative journal questions about their satisfaction with the tasks. All of the pretests, posttests, and journaling were completed online.

### *Results*

The first of two 2-way RM-ANOVA's was used to analyze the ASE data. Subjects' average ASE was calculated from the 23 individual items from both the pretest and posttest.

With group and gender as between-subjects factors and time as the within subjects factor, no significant within-subjects main-effects were found ( $\eta^2 = .10$ ). In the tests of between-subjects effects, there was a significant main effect for group, and analysis of the means indicates that, on average, the experimental group had higher reported ASE than the control group. While the interaction between time and group was not found to be significant ( $\eta^2 = .10$ ), it is interesting to note that it would be significant at the 85% confidence interval and analysis of the mean scores across time for both groups show the control group's ASE declining while the experimental group's increased slightly. Again, these results are not significant but raise issues for future research.

The second RM-ANOVA investigated the mean effects for treatment and gender in the ODA of the chemistry concepts covered in the CAI module. These data were analyzed with  $\eta^2 = .10$  and no statistical significance was found. Naturally, there was a significant main-effect for time as a within-subject comparison, but given the nature of the study, this is to be expected. Also, there was a significant between-subjects effect for the interaction of gender and group. This result reveals that, on average across the pretest and posttest, that females in the experimental group had a higher mean than the control group, while for males, the control group had the higher mean than the experimental group. Again, this is descriptive information, but it does not help the researcher validate or reject any posed hypotheses. Descriptive data are summarized in tables 3-4. Summaries of the RM-ANOVA's can be found in tables 5-6.

### *Discussion*

The findings of the RM-ANOVAs did not support any of the proposed hypotheses. The first, that there will be a significant difference between the two treatment groups on both measures, was not found in either RM-ANOVAs at the 90% confidence interval. This indicates

that, for this small study, there was no significant difference in the effects of receiving a constructivist vs. an objectivist-type feedback approach in a CAI learning module on either ASE or performance on an ODA. However, the interaction of group by time in the ASE analysis warrants further research to investigate the possibility that the type of feedback does play a role in the development of ASE. Furthermore, since ASE is directly related to academic performance, a more refined study may shed light on resultant significant gains as the result of increased ASE.

The second hypothesis, that females would have a higher initial ASE than males was not found to be significant. However, observing the means for the two genders show females slightly higher than males on overall ASE. Finally, the hypothesis that both groups would show an increase in ASE over time was only partially supported. The interaction between group and time reveals that the groups were different across time with the experimental group raising overall ASE; however, the reverse effect was seen in the control group. Time, itself, was not a significant within-subjects main effect indicating overall, that the levels of ASE in an individual did not change significantly between the pretest and the posttest.

**Table 3:** Descriptive statistics for the measure of academic self-efficacy

	Gender	Group	Mean	Std. Deviation	N
AVGSE1	Female	Control	81.57	5.99	11
		Experimental	84.39	10.39	11
		Total	82.98	8.41	22
	Male	Control	80.79	9.93	11
		Experimental	82.96	4.71	11
		Total	81.88	7.66	22
	Total	Control	81.18	8.01	22
		Experimental	83.68	7.91	22
		Total	82.43	7.97	44
AVGSE2	Female	Control	80.53	6.76	11

	Experimental	86.08	9.22	11
	Total	83.30	8.38	22
Male	Control	77.36	10.03	11
	Experimental	83.27	5.56	11
	Total	80.32	8.47	22
Total	Control	78.94	8.50	22
	Experimental	84.67	7.57	22
	Total	81.81	8.47	44

**Table 4:** Descriptive statistics for the objective driven assessment

	Gender	Group	Mean	Std. Deviation	N
PERCODA1	Female	Control	29.67	11.83	11
		Experimental	34.93	13.98	11
		Total	32.30	12.92	22
	Male	Control	31.10	11.15	11
		Experimental	28.71	10.62	11
		Total	29.90	10.70	22
	Total	Control	30.38	11.25	22
		Experimental	31.82	12.53	22
		Total	31.10	11.79	44
PERCODA2	Female	Control	34.45	11.62	11
		Experimental	40.19	19.73	11
		Total	37.32	16.07	22
	Male	Control	52.15	26.47	11
		Experimental	31.58	15.96	11
		Total	41.87	23.79	22
	Total	Control	43.30	21.91	22
		Experimental	35.89	18.06	22
		Total	39.59	20.19	44

**Table 5:** Repeated measures ANOVA summary table for the measure of academic self-efficacy

SV	SS	df	MS	F	p
Sex (S)	92.13	1	92.13	0.865	0.358
Group (G)	372.28	1	372.28	3.50	0.069
S x G	0.114	1	0.114	0.001	0.974
s:SG	4258	40	106.45		
Time (T)	8.45	1	8.45	0.334	0.567
TxS	19.43	1	19.43	0.768	0.386
TxG	57.42	1	57.42	2.27	0.140
TxSxG	1.35	1	1.35	0.53	0.818
s:TSG	1012.18	40	25.31		

**Table 6:** Repeated measures ANOVA summary table for the objective-driven assessment of chemistry concepts from the CAI module

SV	SS	df	MS	F	p
Sex (S)	25.50	1	25.50	0.085	0.772
Group (G)	196.74	1	196.74	0.659	0.422
S x G	1586.82	1	1586.82	5.32	0.026
s:SG	11941.58	40	298.54		
Time (T)	1586.82	1	1586.82	7.396	0.010
TxS	264.73	1	264.73	1.23	0.273
TxG	430.94	1	430.94	2.01	0.164
TxSxG	478.78	1	478.78	2.23	0.143
s:TSG	8582.22	40	214.56		

While this study has a very strong design, there are many threats to both internal and external validity. Furthermore, limitations in the form of weaknesses in the CAI module may have played a role in the results of the RM-ANOVAs reported. The primary sources of internal validity threats lie due to mortality. The loss of subjects due to data loss may have had a profound impact on the results that can not be investigated without further studies to see if the results found here are replicable. Threats to external validity include (a) pretest sensitization, (b) Hawthorne effect, (c) novelty effect, (d) experimenter effect, (e) ambiguous independent variable, (f) unrepresentative sample, and (g) nongeneralizable dependent variable. Pretest sensitization is always a concern when the design includes a pretest. Additionally, since the subjects were not only told that they were part of a research study, but were monitored throughout the entire process, the Hawthorne effect may play a role in skewing the reported results. All the students for this study were from one teacher's class; therefore, it is possible that this study repeated with another experimenter in charge and with different students would not reproduce the results. Other researchers may interpret the two levels of feedback differently than this study used. Therefore, feedback is an ambiguous independent variable and a threat to

external validity. Finally, ASE is a construct that no set measure exists for that is widely reputable. Therefore, it is unsure as to whether the ASE measure used in this study truly represents the construct it purportedly measures. While investigations of the measure's internal validity and reliability were performed, further research of this measure is needed to remove this threat to external validity.

Furthermore, the CAI module needs revisions before future use. After reading through selected journal entries for the participants, it is clear that many students had troubles focusing through the entire CAI experience. For the next study, the CAI should be designed to cover less material at once and provide more "check your understanding" options for the learner. Journal comments from the control group also often expressed frustration concerning the lack of feedback on how to solve the mathematics problems in the "check your understanding" sections. This comment was not present in the experimental group's journaling. Without a qualitative study of these entries, however, further conclusions should not be drawn about the learner's preference for the type of feedback.

Future research, therefore, should consider revising the CAI modules to cover less overall content and perhaps include a series of modules where the ASE and ODA performance was tracked throughout the time series. Also, more effort to reduce the mortality of the sample should be made to ensure that the data loss is not significantly impacting the results reported. A qualitative analysis of this study's journal entries should be performed in advance to designing a new trial so as to take into consideration the comments of this trial's participants. Based on the background literature surrounding ASE and performance, the research question still stands as to whether or not feedback in CAI environments can positively affect ASE, thereby positively influencing academic performance. Implications for future research could impact online and

computer-based course design and has the possibility to change the current trend in computer-based assessment that provides little or no constructivist-type feedback to learners.

## References

- Albion, P. R. (2001). Some factors in the development of self-efficacy beliefs for computer use among teacher education students. *Journal of Technology and Teacher Education*, 9(3), 321-347.
- Archer-Kath, J., Johnson, D. W., & Johnson, R. T. (1994). Individual versus group feedback in cooperative groups. *Journal of Social Psychology*, 134(5), 681 - 695.
- Bandura, A. (1994). Self-Efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York: Academic Press.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.
- Bandura, A. (2001, March 2001). *Guide for constructing self-efficacy scales*. Retrieved July 3, 2002, 2002, from <http://www.emory.edu/EDUCATION/mfp/Bandura/Index.html#guide>
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development*, 72(1), 187-206.
- Bong, M. (2002). Predictive utility of subject-, task-, and problem-specific self-efficacy judgments for immediate and delayed academic performances. *Journal of Experimental Education*, 70(2), 133(130).
- Brooks, J. G., & Brooks, M. G. (1993). *In search of understanding: the case for constructivist classrooms*. Alexandria, VA: Association for supervision and curriculum development.
- Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328-375). New York: Macmillan.
- Cassidy, S. (2000). Learning style, academic belief system, self-report student proficiency, and academic achievement in higher education. *Educational Psychology*, 20(3), 307-323.

- Filby, N. N., & Cahen, L. S. (1985). Teacher accessibility and student attention. In C. W. Fisher & D. C. Berliner (Eds.), *Perspectives on instructional time* (pp. 203-214). New York: Longman.
- Gagne, R. M. (1984). Learning outcomes and their effects: Useful categories of human performance. *American Psychologist*, *39*(4), 377-385.
- Graham, S. (1994). Motivation in African Americans. *Review of Educational Research*, *64*, 55-117.
- Griffin, B. W., & Griffin, M. M. (1997). The effects of reciprocal peer tutoring on graduate students' achievement, test anxiety, and academic self-efficacy. *The Journal of Experimental Education*, *65*, 197-209.
- Jackson, J. W. (2002). Enhancing self-efficacy and learning performance. *Journal of Experimental Education*, *70*(3), 243-254.
- Joo, Y., Bong, M., & Choi, H. (2000). Self-efficacy for self-regulated learning, academic self-efficacy, and internet self-efficacy in web-based instruction. *Educational Technology Research and Development*, *48*(2), 5-17.
- Lepper, M. R., & Chabay, R. W. (1988). Socializing intelligent tutors: bringing empathy to the computer tutor. In H. Mandl & A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems* (pp. 242-257). New York: Springer-Verlag.
- Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *The School Psychology Review*, *31*(3), 313-327.
- Madorin, S., & Iwasiw, C. (1999). The effects of computer-assisted instruction on the self-efficacy of baccalaureate nursing students. *Journal of Nursing Education*, *38*(6), 282-285.

- Marsh, H. W. (1992). Content specificity of relations between academic achievement and academic self-concept. *Journal of Educational Psychology, 84*, 35-42.
- Marsh, H. W., Walker, R., & Debus, R. (1991). Subject-specific components of academic self-concept and self-efficacy. *Contemporary Educational Psychology, 16*, 331-345.
- Meece, J. L. (1991). The classroom context and students' motivational goals. In P. R. Pintrich (Ed.), *Advances in motivation and achievement: Vol. 7* (pp. 261-285). Greenwich, CT: JAI Press.
- Miller, J. W. (2000). Exploring the source of self-regulated learning: The influence of internal and external comparisons. *Journal of Instructional Psychology, 27*(1), 47-52.
- Okech, A. P., & Harrington, R. (2002). The relationships among black consciousness, self-esteem, and academic self-efficacy in African-American men. *Journal of Psychology, 136*(2), 214-225.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578.
- Pajares, F. (2000). First Person: Frank Pajares on nurturing academic confidence. *Emory Report, 52*(21).
- Pajares, F. (2001). *They did not give up*. Retrieved October 31, 2002, from <http://www.emory.edu/EDUCATION/mfp/efficacynotgiveup.html>
- Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into Practice, 41*(2), 116-125.
- Pajares, F., Hartley, J., & Valiante, G. (2001). Response format in writing self-efficacy assessment: Greater discrimination increases prediction. *Measurement and Evaluation in Counseling and Development, 33*, 214-221.

- Pajares, F., & Miller, M. D. (1994). The role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Counseling Psychology, 86*, 193-203.
- Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics performance: The need for specificity of assessment. *Journal of Counseling Psychology, 42*, 190-198.
- Pajares, F., & Miller, M. D. (1997). Mathematics self-efficacy and mathematical problem-solving: Implications of using varying forms of assessment. *Journal of Experimental Education, 65*(3), 213-228.
- Pajares, F., & Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology, 24*, 390-405.
- Russell, M. (1999). Testing on computers: a follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives, 7*(20), 45.
- Schunk, D. H., & Pajares, F. (2001). The development of academic self-efficacy. In J. Eccles (Ed.), *Development of academic motivation*. San Diego: Academic Press.
- Smith, S. M. (2001). The four sources of influence on computer self-efficacy. *Delta Pi Epsilon Journal, 43*(1), 27-39.
- Vrugt, A. J., Langereis, M. P., & Hoogstraten, J. (1997). Academic self-efficacy and malleability of relevant capabilities as predictors of exam performance. *Journal of Experimental Education, 66*(1), 61 - 73.
- Wigfield, A., Eccles, J., & Pintrich, P. R. (1996). Development between the ages of 11 and 25. In R. C. Calfee (Ed.), *Handbook of educational psychology* (pp. 148-185). New York: Macmillan.

- Wilson, B. G., Cole, P., Myers, K., & Clariana, R. B. (1994). *Feedback as an instructional strategy: a constructivist perspective*. Paper presented at the American Educational Research Association, New Orleans, LA.
- Wood, R. E., & Locke, E. A. (1987). The relation of self-efficacy and grade goals to academic performance. *Educational and Psychological Measurement*, 47, 1013-1024.
- Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 202-231). New York: Cambridge University Press.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64-70.

## Appendix

The following are screen shots of the CAI module during the “check your understanding” sections of the control and experimental modules.

Question: a sample multiple-choice question that is identical between both groups

The screenshot shows a software interface for a chemistry lesson. At the top, a dark blue banner contains the text "Colligative Properties of Solutions" in white. Below this, a lighter blue banner reads "Concept Check #3: Molality and Mole Fraction". The main content area is enclosed in a rounded blue border and contains a question on the left and four multiple-choice options on the right. A red button labeled "View a Periodic Table Online" is positioned to the right of the options. At the bottom right of the interface is a button labeled "Return to the lesson".

**Colligative Properties of Solutions**  
*Concept Check #3: Molality and Mole Fraction*

**Question:**  
How many kilograms of water must be added to 9.0 g of oxalic acid ( $H_2C_2O_4$ ) to prepare a 0.25m solution?

0.0025 kg  
 0.025 kg  
 0.40 kg  
 4.0 kg

[View a Periodic Table Online](#)

[Return to the lesson](#)

Experimental Group: feedback for an incorrect answer

Incorrect.

Possible sources of error:  
Incorrect mathematics or formula,  
type of compound, misunderstanding  
of concept

Control Group: feedback for one particular wrong answer (feedback changes depending on the answer chosen)

There's a problem with your unit cancellation...

Can you find it?



Remember, molality is moles of solute per kg of solvent....

Your solute, in this case, is oxalic acid and the solvent is water...



So, start with 9.0 grams of oxalic acid. Your next conversion should have *g* of oxalic acid on the bottom and moles of oxalic acid on the top...

(hint... use the molar mass!)



Your next conversion should involve the molality... make sure the moles of oxalic acid cancel...

Try again...