The Dissertation Committee for Deborah Rush Walker

certifies that this is the approved version of the following dissertation:

INVESTIGATION OF FEEDBACK ON STUDENT PERFORMANCE

Committee:

____
Joseph J. Lagowski, Supervisor

____
Marilla Svinicki

____
Jennifer Brodbelt

____
John M. White

____
Alex Kilpatrick

INVESTIGATION OF FEEDBACK ON STUDENT PERFORMANCE

by

Deborah Rush Walker, B.A., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May, 2005

Dedication

To my deceased paternal grandparents:

Leslie, grandfather, for always asking, “What have you learned in school?”

then really listening (and quizzing!) my responses

Maggie, grandmother, for being the source of courage,

exuberance, and persistence

Acknowledgements

I am so very blessed to have a great many people accompanying me each step of the journey we call life. Here I list a portion of the colleagues, family, and friends who have contributed to the completion of this work. Many heartfelt thanks to:

Joe Lagowski, my supervisor, for catchy phrases like “It is what it is”, helpful feedback, and for untiring exploration of ideas.

The Chemical Education research group at UT for feedback, engaging conversations and delightful chats covering all manner of topics.

The Department of Chemistry and Biochemistry for the housing and resources both to do this work and to share it with others. And the Library system at UT, for a quiet, sealed room in which to think aloud and write without interruption.

To my committee members J. Brodbelt, J. White, M. Svinicki, and A. Kilpatrick for unfailing support before, during, and after the research study was conducted. I feel I benefited as much or more from off-topic discussions as the on-task ones. Benefiting from your years of experience, our discussions guided not only this research but also my personal aspirations.

To Marilla Svinicki, whose encouraging and supportive style cannot help but inspire all who know her. A brilliant researcher and gifted teacher, you embody greatness and have no reservation against sharing it with others.

Alex Kilpatrick, gifted colleague and friend, who generously created the online database on his own time and to my specifications, then maintained it for the duration of the study. Thanks also for the many helpful discussions during the design of the study.

Paige Hill, for generously creating the graphical interface we used for the quizzing and feedback system.

Dr. White, for allowing me to conduct research in your classroom, for allowing me the freedom to stick to my methodology even when it meant more work for you, and for allowing me to invade your classroom time to give surveys. I deeply appreciated and benefited from classroom observations of your lecture style—and your famous guitar solo in class.

Fatima Fakhreddine, who continuously inspires me to be and do more, and whose teaching style inspired this research.

Vickie Williamson, mentor and friend. Always there to listen and to guide me toward the answers I seek, you live what it means to be a great advisor.

Diana Mason, it seems that when I’m frantic, I can email you. Dubbed “sleepless one”, you have always responded to my email queries within a few hours, with information rather than advice. You made delightful my conversion from chemical research to chemical education research. And I won’t soon forget Tihuana, Mexico—it goes to show that some of the best conversations take place outside the office. (This is for you, too, V. Williamson.)

Victoria Stolarski and Debbie Hess, my mentors, friends, and cheering squad who never ceased in believing in me.

Vaughn Nelson, Treasure Brasher, and Gene Carlisle, three teachers from my undergraduate alma mater: West Texas A&M. If I have blossomed, it’s because these dear people found and planted the seed, balancing care, challenge, and excitement for science.

Jerry and Anita Rush, my parents. Much to your credit, you balanced pushing and encouraging me in academics. With daddy a carpenter, I couldn’t help but pick up notions of function and form. With mom the owner of a crafts business, I received ample opportunity to think out of the box and merge seemingly disparate items into something new. I continue to benefit from these and other skills.

Marcus Walker, my husband and true love, without whom this work would not have been possible. Not only did you code the spreadsheet software to do the calculations I had planned and assist in formatting this document, during my tenure you also helped me to recover from three hard-drive crashes and four miscarriages, even saving my life when one miscarriage went awry. Just as saying “I love you” doesn’t capture the depth of my feelings, saying “thank you” pales in comparison to all you have done.

Joshua, our 3 year old son, you are a source of incredible inspiration and humility. You help me see common things through new eyes; play time with you is a welcomed respite from the rigors of analysis and writing.

My extended family, including Angela Ritter (sister) and attendees of the annual Rush Reunion, thank you for years of encouragement, familial support, and genuine attempts to figure out just why I’m still in school. Soon I’ll not be, so I’ll have time to practice for the Cow Chip Throwing contest. Y’all watch out, because here I come!

INVESTIGATION OF FEEDBACK ON STUDENT PERFORMANCE

Publication No. __________

Deborah Rush Walker, Ph.D.

The University of Texas at Austin, 2005

Supervisor: Joseph J. Lagowski

Comparison of two types of cognitive feedback differing in specificity forms the basis of this study: one type presents the explanation of the problem (informational feedback) and the other presents the explanation plus a corrective statement identifying what the student did wrong (diagnostic feedback). To do this, we inserted an online quizzing and feedback system into an existing second-semester general chemistry course at a large southwestern university. A sample consisted of 95 college students deriving from one of the large lecture course sections, was randomly assigned to one of the two feedback conditions. The students participated in e-quizzes for the entire semester, receiving one of two feedback types after each question. Feedback type (specificity) and feedback count (number of “treatments”) comprised the independent variables while performance, motivation, and epistemological beliefs comprised the dependent variables. Analysis spanned 154 multiple-choice questions. Results show no significant difference between the two feedback types employed, but scatter plots and participation plots revealed interesting trends. Our results suggest that increasing specificity of feedback in the manner we employed and using a sample and environment similar to ours does not correspond with increased performance and does not correspond with significant changes in motivation or epistemological beliefs when compared to informational feedback that provides a detailed explanation.

Contents

List of Tables

List of Figures

Chapter 1: Introduction

> "…freshman chemistry (general and organic) is probably the most problematic science discipline taught (as far as learning difficulties and misunderstandings are concerned)."
> Uri Zoller, 1990

This research stems from the desire to help students in freshman chemistry. Bloom (1968) recommended diagnosis, feedback and remediation in order to help students. Over time this became drill and practice. Now, with our electronic age, more assignments are converting to the digital and online environments. With each new advance in online tutors, feedback remains a staple. This begs the question: how much feedback is enough? How much is too much? This study cannot yet answer this question completely, but we hope that it can help advance current research in this area.

Students have a difficult time taking chemistry. Doing chemistry problems is known to help students, but only when the students actually learn from the experience. Electronic homework and quizzing has become a popular method of drilling students and many systems are on the market as CDs or websites. But designers building these systems face the burning question of how much feedback to provide. Writing too much feedback is costly, since the text must be designed for instruction, written, proofread, coded, and tested. But writing insufficient feedback is costly to the learner. So where is the balance? This research aims to build upon current feedback research to help further the search for that balance.

Students possess a variety of characteristics that influence their willingness to accept feedback and their desires to improve from it. Feedback assistance can take different forms, from a response of correct/incorrect to individualized feedback tuned to a specific response. Research in feedback shows contradicting conclusions about whether increasing specificity of feedback increases performance and motivation. It is the investigation of specificity of feedback that forms the basis of this study.

We have chosen to investigate the specificity question by creating online quizzes employing two types of feedback: informational (explanation and/or solution of chemistry problem) and diagnostic (statement of error performed plus explanation or solution of chemistry problem). Our study involves freshman chemistry students in a large lecture (N>200) course at a large southwestern university during the spring offering of second-semester general chemistry. The design of this study is based on two goals. One, we wish to explore electronic quizzing in a classroom setting as an option for helping students learn chemistry. Two, we wish to determine which of two levels of feedback specificity help students more and investigate the relationships among feedback specificity, performance, motivation, and epistemological beliefs about knowledge in chemistry. We wish to find out if “more is better” with regard to feedback for general chemistry and our students. We wish to determine whether feedback itself can influence students’ motivational goals with regard to learning, perhaps resulting in students increasing their focus toward mastery of chemistry. And we wish to determine whether feedback can influence students’ beliefs about how knowledge in chemistry is constructed; in particular, to determine whether specificity of feedback influences maturity in chemistry students, with regard to beliefs about knowledge.

Thus, our overarching research question is:

> Does informative or diagnostic feedback have the greater positive influence on performance, motivation, and epistemological beliefs toward chemistry?

We use descriptive and inferential statistics to determine changes and relationships among the variables involving performance, motivation, and epistemological beliefs of chemistry.

Limitations

With regard to feedback and selecting experimental groups for our study, we accept the results of former research that establishes that the no-feedback case is inferior and we turn our focus to the comparison of two feedback types with one another.

In that focus, the definition and implementation of feedback types are somewhat open to interpretation and so we wish to make clear our definitions of each. Informational feedback provides an explanation or solution to the assignment question. In contrast, our use of diagnostic feedback first seeks to confront the error the student likely made in choosing the incorrect response, then to correct it with an explanation of the assignment question. Thus, we will refer to this as a “confront and correct” approach to diagnostic feedback. This approach contrasts with others’ (e.g. Butler and Nisan, 1986) applications of goal-based feedback that informs the students of their proximity to meeting the objectives of the assignment.

Another limitation of this study is that this is a field study, designed to be incorporated into an existing classroom with minimal change from the *modus operandi* of the course. While e-quiz format was specifically chosen in order to communicate a more controlled attempt of “post-studying” assessment than that of homework, the students completed the quizzes online and therefore may or may not have studied the topic beforehand. The students were asked not to use textbooks or peers for help, but this aspect could not be controlled. Therefore, conclusions from this study must be received with the understanding that the study was not conducted in a laboratory, or controlled, setting.

The nature of the online quizzes limited the interactions between student and computer since the e-quizzes were conducted as part of a “normal” course and functioned much like a “normal” quiz. Thus, the limit was one interaction (feedback intervention) per question. While the point was to determine which feedback to use in the single-interaction example, it limits the comparison of results to other studies that allow multiple interactions (e.g. multiple attempts) during a learning event.

The nature of the feedback places an important limit on the interpretation of feedback results. The feedback differed only for corrective feedback, i.e. when the student provided an incorrect response. Therefore, the “treatment” occurred when the student answered in error and this, obviously, did not occur after every question. In addition, since the e-quizzes were voluntary and offered only as extra points for a limited group of students, participation was inconsistent across the study term so results are skewed toward students who opted for doing the assignments.

Finally, the measures of motivation goal orientation and epistemological beliefs were taken from survey responses. The epistemological beliefs survey is still being perfected. That is, researchers are continuously improving the paper-and-pencil survey to capture epistemological beliefs.

Overview

We begin with chapter 2, a literature review of the three areas under study, followed by the research questions of the study. Chapter 3 incorporates the findings from the literature into a methodology for studying the hypotheses, then chapter 4 presents the results of those tests and follow-up analyses. Chapter 5 leads the reader through discussion of the results, implications, and areas for further research. Surveys, questionnaires, and forms are provided in the appendices.

Chapter 2: Literature Review

During office hours an instructor has the opportunity to offer very specific feedback attuned to a particular student's needs. During this one-on-one time the teacher examines the individual student's work and questions the student in order to discover both what the student knows and where the student strayed. A dialogue follows in which the teacher tries to help the student realize and acknowledge his error. The instructor may then gently correct the student's misconception or miscalculation and assist him through the correct steps. The teacher in this scenario has diagnosed the student's error(s) and assisted the student.

This student-teacher interaction is typically limited to office hours, perhaps because of time constraints, the student's fear of looking stupid, or lack of interest. With a computer-based feedback system however, the issue of time constraint is diminished, as is the fear of looking stupid, since the interaction is with the computer rather than the professor. Additionally, by providing feedback that is immediately available within the quiz, lack of interest can be circumvented. This then begs the question: what feedback should be used to mimic the outcomes in the above scenario?

Gagne's "Nine Events of Instruction" echo a similar prescription for what he referred to as 'effective learning', which includes the necessity of feedback (1965, 1992). In 1968, Bloom supplied a straightforward prescription for helping students learn: diagnosis, feedback, and remediation (1968). Following similar thinking, other theories of learning include such components as presenting the material, testing, and providing feedback.

Feedback and Performance

Yeany, et al, in two studies and a meta-study (1979, 1980, 1983) reported that while Bloom's (1968) prescriptions had good intentions, remediation had no effect beyond that of diagnosis and feedback alone. Diagnosis was accomplished in the formative testing of the students, followed by either feedback alone (referred by the researchers as diagnostic feedback) or feedback and remediation. They reported: "The surprising result of this study is the source of the impact [on performance]: It does not appear to be the remediation but rather the diagnostic feedback." We thus determined that feedback appropriately given might be adequate for helping students learn a subject.

A continuing survey of the feedback literature produced an immense number of articles spanning research areas, methodologies, and content topics, but falling into two general categories that mirror two primary educational theories: behavioral (motor skills) and social-cognitive

(cognitive skills and content knowledge). While the information in the area of motor skills research is interesting and will be referred to on occasion, this review focuses on the topics of concept learning or information learning, as they better describe methods for improving academic performance. But before we continue with a review of feedback literature, we must become familiar with a few terms.

A categorization scheme by Holding (in Bilodeau, 1969, p.257) aids in the classification of feedback studies and future research. The scheme begins with two divisions for feedback: intrinsic and artificial. Intrinsic feedback is internal to the task while artificial feedback comes from a third-party source or a source external to the task and the participant. For example, when one tosses a free throw in basketball, one can know whether the ball is going into the hoop at the instant the ball leaves the hand, even before it reaches the hoop. This is an example of intrinsic feedback. A chemistry example might be a student who performs a density calculation to get an answer of 5,000 g/mL and instantly suspects that something is wrong. In both cases, each received feedback from internal sources (muscular, physiologic internal feedback for the basketball player; cognitive comparisons for the chemistry student). On the other hand, artificial feedback comes from an external source, like a coach or a teacher. Intrinsic feedback is inherently difficult to capture or measure, while artificial feedback is easily controlled. This study manipulates artificial feedback.

Further classifications of feedback presented by Holding (in Bilodeau, 1969) provide a basis for discussion and categorization of current and future feedback studies. These terms include: concurrent or terminal, immediate or delayed, nonverbal or verbal, and separate or accumulated. Concurrent feedback is given during the task, whereas terminal feedback is given after completion of the task. The next level consists of immediate and delayed feedback. This one can be tricky because the time “cut-off” between immediate and delayed feedback is hazy and differs with the reference of the paradigm with which the study is associated, either motor skills or cognitive skills/content knowledge. Motor skills research considers 10 seconds as delayed feedback while that same time interval is considered immediate feedback in the cognitive skills or content knowledge paradigm. For the next comparison, nonverbal feedback comes from a machine, like the gauge for the gas tank on an automobile, as opposed to verbal feedback that includes any feedback consisting of words, be it audible or not (including text). The classifications between separate and accumulated feedback relates to the pairing of feedback to task. If the feedback is given after each task, then it is separate. Accumulated feedback indicates the case in which several tasks are conducted before feedback is given (as for a typical examination).

This study focuses on feedback given to individual online quiz questions immediately after student response. Thus, according to Holding's classifications, the feedback in this study is categorized as artificial (external to student and task), terminal (after the task, not during), immediate, verbal (text), and separate (after each question/response). This information is provided to help the reader understand how this study fits within the vast feedback literature. However, Holding's classification terms are limiting for extensive use here since they do not address the central question of this study. Another classification scheme is necessary and will be presented shortly.

The feedback literature is rich and expansive, including a multitude of topics over the last hundred years. Ammons' review of the literature is one among many that reveals that some feedback leads to more benefits than no feedback (1956; for others, look into the Knowledge of Results, KR, research including Judd, 1905-6; Pressey, 1950; Annett, 1969). In chemistry instruction, Angell demonstrated that some feedback was better than none for his chemistry examinations (1949). Angell used a derivative of Pressey's punch-board for testing, in which the student would punch through the top sheet at the spot corresponding to her answer, revealing feedback in the form of color (red=wrong answer) as she took the examination. Similar results were found by Hall, et al, (2001) for homework assignments. Further, the same review of the literature by Ammons states that more specific feedback yielded increased rates of learning, increased level reached by the learner, and increased motivation (1956). He also warns that in some cases feedback may impede learning, citing learner's detrimental dependence on the feedback in one case. Thus, Ammons proposed that perhaps there exists both an optimum amount of feedback and a maximum amount of feedback, above which no more improvement is found.

The early studies share the drives of present-day studies—those merging pragmatic and theoretical categorizations. Historically, feedback in general seems to have a pragmatic feel to it, but two 20th century papers in motor skills research take a somewhat theoretical approach and offer valuable guidance to the setup of our study on specificity. Thorndike (1927) wanted to know if outcome feedback could affect subsequent behavior. His results supported this hypothesis. Trowbridge and Cason (1932) conducted a follow-up study that replicated Thorndike's study in many ways, but also expanded it. The Trowbridge and Cason experiment found that the effects of feedback on scores depends on the amount of useable information in the feedback and found that all feedback can influence performance, but not all feedback is helpful. A nonsense feedback trial had lower performance than the no feedback trial.

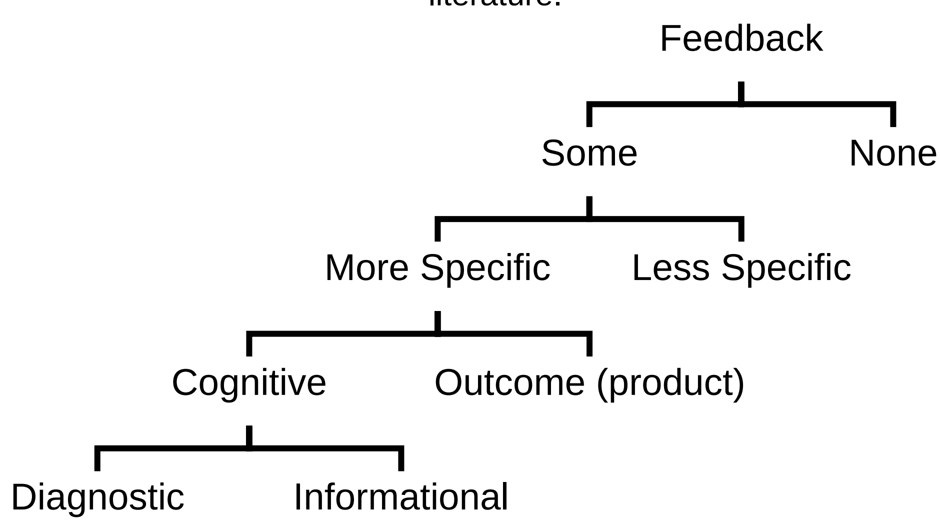
Finally, two more points are worth mentioning. For one, the Trowbridge and Cason study supported Thorndike's position that feedback can cause changes in subsequent performance and they suggested that it was the information—the kind and specificity of information—that led to changes in performance. Research in the "kind" of information in feedback has been studied, but research in specificity is still lacking.

Feedback and Specificity

A change in specificity of the information in the feedback provided was noticed among studies getting varied results. So we began a hierarchical categorization of feedback studies based on specificity of feedback. Figure 1.1 represents a hierarchy we created when categorizing feedback studies in content knowledge and cognitive skills, this time based on specificity of the feedback employed.

The first studies concentrated on whether recipients of some feedback outperformed those receiving no feedback (see, for example, Thorndike, 1927; Annett, 1969). Indeed they did in all cases except for nonsense feedback (Trowbridge, 1932). Looking at the great many feedback studies focused on knowledge of results (KR—notation of right/wrong) or knowledge of correct response (KCR), one wonders whether there exists feedback with the intention to help the learner construct valid arguments about a topic, a sort of cognitive feedback, distinguished and apart from outcome feedback. Indeed, researchers continue to look at increasing the information content of their feedback to employ cognitive-based explanations for their learners—we refer to this type of feedback as cognitive feedback.

Figure 2.1: Classifications of feedback, based on a review of the feedback literature.



Prior Uses of Cognitive Feedback

Cognitive feedback has taken many different forms in the feedback literature: descriptive feedback (Taylor, et al 1984), process feedback (Early, et al, 1990), cognitive feedback (Jacoby, et al, 1984), and learning-oriented feedback (Dweck, 1986). The research of this study will focus on comparing two types of cognitive feedback: informational and diagnostic (to be defined later in this chapter).

For the purposes of this research, the term cognitive feedback is used as a general category in which each of these forms of feedback exist. In essence, cognitive feedback indicates to the learner the what, why, and how a solution is correct for the given question. In contrast, outcome feedback provides merely the statement “correct/incorrect” or normative information. Cognitive feedback tends to yield better performance than outcome feedback (Butler & Nisan, 1986; Jacoby, et al, 1984; Earley, et al, 1990; Johnson, et al, 1993; Moreno, 2004; Goodman, et al, 2004), though the exceptions are still not well understood (Kluger and DeNisi, 1996).

Recent research in feedback has tended to focus on the study of cognitive feedback. Different researchers hold slightly different views of cognitive feedback, seemingly based on the researchers’ philosophical beliefs about the function of feedback, which include: guidance, cueing, sign, incentive, command, informing, or goal realignment (adapted from Bilodeau, 1969). Three such approaches to cognitive feedback are herein discussed: (a) quantity of information, with the thought that more information—elaboration, for example—is better, (b) learning-oriented versus performance-oriented feedback, and (c) goal-based feedback that is intended to realign the student's efforts to the learning goal and help the student redefine that goal in his mind. We now describe examples of each and their possible contributions and limitations with regard to quizzing feedback in chemistry instruction.

Feedback: Quantity of Information

In an example of information quantity, Jacoby et al., (1984) investigated feedback for security analysts' decision making. The researchers compared outcome feedback (referred to as "accuracy of response") and cognitive feedback ("information value"). Outcome feedback was defined as "information that describes the accuracy or correctness of the response" while cognitive feedback "represents information regarding the how and why". The outcomes of the Jacoby study (1984) are similar to Trowbridge and Cason's (1932) as are the comments: both researchers believe that feedback influences performance by its information value, as opposed to influencing performance by affect, which Thorndike originally proposed. Jacoby, et al, further suggests that the value of such feedback may lie in its predictive and explanatory value, suggesting that in order to be effective, feedback needs to offer information that explains the answer and also offers information that provides the learner with prediction-making ability. Take, for example, a chemistry problem in which the student is asked to rank several molecules in order of increasing boiling points. Outcome feedback would only identify whether the student was right or wrong. Explanatory feedback would include a statement explaining the answer to the problem, for example, that molecular hydrogen has a lower boiling point than ammonia because the latter has intermolecular forces, such as hydrogen bonding, acting among the molecules to increase the boiling point over that expected based on molecular weight. The predictive aspect of feedback can occur alongside explanatory feedback, but does not have to. In this example, informing the student that intermolecular forces and hydrogen bonding is key to solving this problem allows the student some aid in performing future problems of this type. Predictive value may be increased with feedback that states, for example, that boiling points vary with intermolecular forces and weights, in that order, such that molecules with the strongest intermolecular forces have the highest boiling points and for molecules with the same intermolecular force strength, the heavier molecule has the higher boiling point. Thus, Jacoby et al (1984) suggested that the effect of feedback on performance may rely on both the explanatory and predictive values of the feedback.

Feedback: Learning-oriented versus Performance-oriented

In an example of a process-related versus performance-related feedback study, Johnson, et al, (1993) asked students to operate a Space Shuttle Remote Manipulation System simulation. Two feedback conditions were given: “performance oriented” feedback that was outcome feedback (evaluation of “correct/incorrect”) and “learning-oriented” feedback that consisted of outcome feedback plus a statement of correction. The researchers expected learning-oriented feedback to “enhance performance both through its information value and through its cuing function” and thus lead to improved performance. This prediction was met, as fewer errors occurred under the learning-oriented feedback condition.

Feedback: Goal-related

In an example of goal-related feedback, Butler and Nisan (1986) compare “individualized task-related” feedback against normative (outcome) feedback and against no feedback conditions in their study involving sixth graders studying English grammar over a three-day period. The authors propose that providing “diagnostic information relevant to self assessment” promotes a “sense of mastery and self-determination”. The task-related feedback consisted of a complimentary phrase and a correctional, goal-related phrase. For example, for a task in which students were given a word and asked to make a list of words using those letters, the sample task-related feedback was “You wrote many short words, but not many words”. Since this type of feedback gives more information than the normative feedback (grades), it is herein categorized as cognitive feedback, though its content differs from that in the Jacoby, et al, study (1984). In addition, since the feedback given in Butler and Nisan’s (1986) study gives information about the goal of the task (inferred) and the proximity of the student’s work to that goal, we refer to it as “realign and redefine” cognitive feedback, as it is intended to realign the learner’s efforts to the learning goal and assist the learner in redefining her understanding of the goal. This investigation draws on Butler and Nisan’s work for its design ideas, but differs in subject area, participant age, and duration of the study.

Implications and issues

One can see from these examples that the construction and implementation of cognitive feedback can change slightly in different applications, but the underlying principles remain similar: to identify what, how, why, and how far off (from goals), or for quantity of feedback, provide increasing amounts of information up to stating the correct response and the steps to solution.

While effective in some learning environments, the goal-based "realign and redefine" implementation of cognitive feedback may not carry over to learning of chemistry concepts or calculations in which informing in simple terms "how far off" may not add to one's knowledge. Feedback's usefulness is said to be determined by the addition of new knowledge. To illustrate, imagine the following goal and question pair:

> Goal: Be able to determine the amount of product that may be formed from a given amount of reactant.
> Question:
> Given: 2 NaOH + H_2SO_4 → Na_2SO_4 + 2 H_2O
> How much Na_2SO_4 (in moles) may be formed from the reaction of 3 moles of NaOH with excess H_2SO_4? Assume 100% yield.

If the student answers wrongly and if goal-based feedback were to be given that matched the previous studies discussed, then that feedback would need to give information on the separation between the student's performance and the desired goal. One example would be for the student to answer 3 moles. The goal-based feedback might then state that the student did select a response measured in moles, but did not accurately calculate the number of moles of product formed.

Such feedback would follow the “letter of the law” given in prior studies, but does not fit “the spirit” in other researchers’ use of cognitive-style feedback (e.g. Jacoby, et al.1984, Johnson, et al, 1993; Moreno, 2004; Earley et al, 1990) because it does not inform what, how, or why another response is better so it does little to inform the student how to correct the error. Therefore, another construct must be developed; one that fits within cognitive feedback but is more specific than the generic explanation of the assignment question. In devising this new construct, the literature was again examined.

Toward a New Feedback Type

Kulhavy (1977) discusses that corrective feedback needs to “not only eliminate the wrong answer, but also to substitute correct information in its place”, but does not offer direct instruction on how to go about creating this type of feedback. Some information is provided in Jacoby, et al, (1984) in which the authors recommend that feedback have “predictive power”, a construct better defined in Jonassen, et al, (1993), as “criteria for analyzing the performance” rather than “direct instruction about how to perform certain skills”. Even with these examples from the feedback literature, some question remains as to what these researchers mean by "criteria" and "predictive power" so other resources were consulted, including those in conceptual change research, misconceptions research, and teacher training research.

Focus on Knowledge Base

Taconis, et al, (2001) presents a model of capacities needed for effective problem solving in science: knowledge base and skills base. They offer a comparison of information that enhances a knowledge base (general knowledge) and information that expands or develops knowledge of strategies (ability to perform). While both are needed, Taconis, et al, purports that the knowledge base is more valuable in planning effective treatments. Since the knowledge of strategies is focused more on the "doing" and less on the "analyzing", it follows that the knowledge of strategies would have less predictive value (i.e. information that can be transferred to other situations). Information that adds to the knowledge base, in contrast, is more likely to have predictive power to help the student solve future problems.

Successful Teacher Strategies

Turning to strategies employed by successful teachers in handling students' responses, one paper in particular stood out as particularly helpful in determining what diagnostic feedback for the sciences might look like. Hashweh's (1996) research gives insight into characteristics of feedback, though she does not call it feedback. In her study, teachers were given scenarios consisting of questions and student responses and then were asked how they would score and respond to each student. In analyzing the data, Hashweh formulated a list of response types and grouped them according to strategy employed. Results from that study show that Type I strategies (explanation) were the least effective, whereas a use of multiple strategies was most effective in producing conceptual change. Looking carefully at her list and definitions, one can extract potential delineations between informational and diagnostic feedback types. The strategies that explain, repeat, and convince fit the characteristics of informational feedback as previously discussed while refute and restructure provide a strategy for writing diagnostic feedback. Because of our intention to examine feedback for online quizzes in a field study application, the "develop" strategy is employed in this study only inasmuch as the response chosen by the student reflects a misconception or error held by the student that is confronted by the diagnostic feedback.

Misconceptions in Chemistry

So, an important note in employing this definition of diagnostic feedback is the value of identifying and somehow managing the student's misconception or error. For the case of an online quiz and feedback system, an error must be predicted in advance and added to the multiple-choice options as a distracter so that the diagnostic feedback can be written to confront it and replace it with appropriate explanations. This set of strategies then mimics those used by more successful teachers and also fits the strategies recommended for handling misconceptions in chemistry (e.g. Zoller, 1990; Nakhleh, 1992; Gabel, 1999; Herron, 1996; Taber, 2002) and conceptual change research (Posner, et al, 1982). Within motivation research, Dweck (1986) posits that questions need to be challenging, but maintained within a learning-oriented context.

Characteristics of Effective Feedback

Before implementation, these ideas gathered from related research must be balanced with known characteristics of effective feedback (Ilgen, et al, 1979): positive, from a credible source, relevant, focused on the task rather than on the learner, and specific, such that information was gained above and beyond that which the learner already knew. Thus, a new form of cognitive feedback was developed that incorporated each of these results: diagnostic feedback.

Diagnostic Feedback

Returning to our earlier goal-question example, now the diagnostic feedback for the student who answered 3 moles might be written as

> "3 moles is the amount of water formed, but not the amount of sodium sulfate formed. The mole-to-mole ratio between NaOH and Na_2SO_4 in the balanced chemical equation is not 1-to-1. It is 2-to-1. One mole of Na_2SO_4 is formed for every 2 moles of available NaOH reacting with excess H_2SO_4. This may be shown mathematically as…"

The latter two sentences represent the informational feedback portion while the addition of the first three sentences distinguishes the diagnostic feedback.

The use of diagnostic feedback in this study thus employs lessons learned from prior uses of cognitive feedback, merges that with general qualities of effective feedback, includes a balance between knowledge base and knowledge of strategies, qualities of successful teacher strategies, and lessons learned from misconceptions in chemistry.

Feedback and Motivation

Motivation and feedback literature have been tied together almost since feedback research began and it makes sense. Feedback can have many functions (cueing, incentive, goal-orienting) that reflect motivational constructs; separating these can become a challenge. In fact, early feedback researchers wondered whether feedback and motivation could ever really be separately measured (Annett, 1969). Motivation researchers gave similar response—feedback is mentioned numerous times in Weiner's (1990) summary of the history of motivational research. As mentioned previously, Dweck's research concentrated on motivation but gave detailed information about the effective use of feedback (Dweck, 1986, 1991) and many of the examples in the feedback section of this chapter incorporated motivation in their studies (e.g. Butler and Nisan, 1986). Obviously, motivation is an important aspect of the feedback mechanism and, accordingly, it is measured and monitored in this study, as well.

Looking across the feedback literature, the motivation that was measured took a variety of forms, often being limited to the leading motivational theory of the time. It makes sense that early feedback studies, for example, would take an incentive view or reward/punishment view of motivation and explain feedback through that lens. But in the present day there are many theories of motivation from which to choose, so a study's goals can guide measurement decisions.

The initial goal of this study was to help students now and in the long-term by providing feedback during online quizzes. The idea was to formulate the feedback so that students' errors were corrected in such a way as to increase odds of future success. But imagining even a single feedback statement brings to mind motivational influences, as it would seem that any collection of words could carry some motivating (or de-motivating) aspect to it. Measurement would be key. Looking at prior feedback studies and the various motivational constructs known to researchers today (for an overview, see Pintrich & Schunk, 1996, or Linnenbrink & Pintrich, 2002), we chose achievement goal theory, a motivational construct based not on whether someone has motivation or not, but focuses rather on the nature of the motivation.

Achievement Goal Theory (AGT) includes three aspects of motivation, each of which is viewed as a separate entity. In this theory, students are viewed as having motivation in one or more of these three areas. Thus, it is important to distinguish AGT from common beliefs about motivation: in this theory it is not valid to say a student has no motivation. There is no such thing. Instead, AGT looks at what motivates a student: the goal to be the best, the goal to better oneself, or the goal to avoid embarrassment. The student driven by goals to be the best, to outperform others and to demonstrate or show his skills to others is said to have high performance goal orientation. As shown in Table 2.1, these students tend to value ability and use surface strategies like memorization, a reflection of their view that people with high ability should not have to work hard (expend effort) to do a task and that, in turn, people who must expend effort to do a task must not be of high ability. In contrast is the student possessing mastery goal orientation. Such a student values learning and effort. This student tends to think that added effort results in additional knowledge and is therefore “worth it”, so these students tend to use a variety of study strategies, concentrating more on learning than on the grade. This measure is closest in nature to intrinsic motivation. The most recent addition to the goal orientation theory is the performance-avoid goal orientation. Students high in this motivation are driven by a desire to avoid embarrassment. These students tend to select tasks that they know they can complete easily and successfully, as they tend to believe that failure suggests incompetence and incompetence must be avoided. These students tend to fear failure, so their study strategies are often limited to copying templates provided by an external, reliable source.

Table 2.1: Summary of differences between goal orientations. Adapted from Ames and Archer (1988)

	Mastery Orientation	Performance Orientation	Performance-Avoid
Value placed on	Learning and effort	Ability	Ability
Reason for satisfaction	Working hard, challenge	Doing better than others	Avoiding failure or embarrassment
View of errors	Part of learning	Anxiety eliciting	Anxiety eliciting
Strategies employed	Diverse	Practiced/Memorized	Copied/Templates

Achievement goal theory offers a motivational construct that is more malleable than the intrinsic/extrinsic view of motivation and is not task-dependent like self-efficacy. Additionally, achievement goal theory was chosen for use in this study because it has ties to the feedback literature (e.g. Butler & Nisan, 1986), it has ties to performance in academic settings (e.g. Ames & Archer, 1988), and it has the potential of producing interesting effects with the diagnostic, confrontational style of the feedback we are examining. The Milner-Bolotin instrument we use to capture the motivation measures is a pencil-and-paper survey based on Likert-scale responses, shown to be reliable in a setting much like ours: college-level, science classroom.

With regard to potential effects of motivation on feedback, we anticipate that students possessing a high mastery orientation will be most likely to accept their error, learn from it, and move on more than those students with low mastery orientation. Prior research shows that students with mastery goal orientation (with a focus on learning rather than on grades) are more likely to work through challenges on problems, employing a wider variety of strategies in their solutions. They are more able than their peers to “work through” difficulties, so we wonder whether these students will more readily accept and use the feedback provided, including diagnostic-style feedback. Likewise, since students with a high mastery orientation tend to use deeper processing methods in learning, we anticipate them to be more likely to adapt to the corrective feedback and increase or maintain performance over many questions of varying difficulties. On the other hand, students with a low mastery orientation may be more likely to scan or misinterpret negative, corrective feedback and may not learn from it. With regard to the performance goals measure of achievement goal orientation, we anticipate that students with a high performance orientation will be more likely to accept feedback without questioning it (“cloning”), whereas those students having a low performance orientation coupled with a low mastery orientation may skim the feedback or even not read it at all, so they are expected to be the least likely to pay attention to and incorporate feedback into their knowledge structures and therefore not show performance improvements over time.

We also wonder about potential effects of feedback on motivation. That is, if we assume that feedback is information and, as such, can have an effect on motivation, then we wonder whether these effects could be different for different types of feedback in this study. Diagnostic feedback, by modeling the mastery-oriented study strategies of seeking out why the response was wrong and why the answer is correct, may increase or sustain both performance and mastery orientations. Diagnostic feedback may also increase both mastery and performance oriented motivation by reducing the frustration that accompanies not knowing what went wrong. Additionally, diagnostic feedback might promote mastery goal orientation by providing feedback that focuses on the infraction rather than the student, that offers correction rather than punishment, and provides the what, why and how of the explanation. Lastly, we wonder whether the beneficial aspects of diagnostic feedback might manifest most for complex, higher-level application or analysis problems since those offer the greatest opportunity for a student to become lost and frustrated or for cognitive load to be filled.

For the reasons just described, we focus motivational measures on achievement goal orientation and look at both sides of the argument: at whether motivation influences performance in the acceptance of correctional feedback and whether that correctional feedback influences change in motivation.

Feedback and Epistemological Beliefs

Contrary to popular thought, feedback's effect on performance has been variable. Including motivation in the models has helped resolve some issues, but better models are still needed. For example, we now know that the old adage "more feedback is better" does not necessarily hold true because motivation also plays a role and must be considered. Looking across the vast feedback literature, there are still occasions when more feedback led to higher performance or in which more feedback led to lower performance, even with motivation controlled. The explanation for this effect is still unresolved. We propose that the difference may be in the acceptance of feedback and its incorporation into the student's knowledge structures. The aspects of credibility and source of feedback are important to the acceptance of feedback (Ilgen, et al, 1979) such that if a student receives corrective feedback from a non-credible source, that feedback may be quickly discarded from memory and change of behavior or knowledge structure does not occur (Ilgen, et al, 1979). In similar research, Posner and Strike (1982) state that believability of the alternate viewpoint is important for initiating conceptual change. Motivation constructs do not capture these aspects and performance captures only the outcome. Beliefs about knowledge, how it is formed and where it comes from may give some insight.

Research in student's epistemological beliefs offers measures for the constructs of believability and credibility in a cognitive context. Epistemological beliefs have been tied to performance on academic tasks (Schommer, 1990, 1993; Wilkinson & Maxwell, 1991). In this research study, a student's beliefs about whether authority holds the truth are used to measure the likelihood of a given student to accept the corrective feedback provided. The student's beliefs regarding how simplistic and certain is the subject of chemistry measures the likelihood of that student to concentrate on the many complicated steps required to do the chemistry problems presented in this research.

A student's epistemological beliefs are currently captured either by interview or by pencil-and-paper survey. While interviews are much broader in scope, surveys have been shown to be reliable when compared with interviews (Hofer, 1997). There are many different constructs within epistemological beliefs literature, with the choice depending on what one wants to measure. Since our interests involve academic outcomes in a classroom setting, we chose the "Schommer" line of research (Duell &

Schommer-Aikins, 2001; Hofer, 1997; Qian & Alvermann, 1995; Schommer, 1993).

The original pencil-and-paper surveys created and tested by Schommer (1990, 1992, 1993) consisted of more than sixty questions and involved some motivation-sounding subscales. Qian and Alvermann (1995) refined Schommer's original survey, taking the statements that best predicted performance and establishing a more elegant factor structure. Hofer then followed with her deep study across the major epistemological beliefs constructs and created a survey that spanned the various beliefs and that compared nicely with the traditional interviews. The measures that Hofer's survey captures include the simple/certain dimension, the belief in authority dimension, and the attainability of truth dimension. The latter dimension is new to paper-and-pencil surveys in epistemological beliefs, but did predict performance. The "Attainability of Truth" dimension attempts to measure one's agreement with the possibility that truth can, in fact, be found in a given subject area. Finding expression in the Reflective Judgment Theory (RJT) interviews (King & Kitchener, 1974), this construct attempts to better capture ideas about truth that originated in Perry's (1968) seminal work. The student who possesses a high belief in the attainability of truth is said to have a naïve belief about how knowledge comes to be. Hofer's simple/certain dimension merges two separate Schommer dimensions: belief that knowledge is simple, not complex and belief that knowledge is certain, fixed, and unchanging. The "Authority" dimension aims to measure one's dependence on authority for answers, or one's trust in authority over other sources. One who has a high score in "Authority" perceives textbooks, professors, and other authorities as being sources of truth, even if that student's experiences give evidence to the contrary. This measure was chosen for this study because of authority's role in acceptance of feedback. Ilgen, et al, (1984) among others, noted that recipients of feedback are less likely to accept and act upon feedback received from non-credible sources and noted that expertise and trustworthiness factors comprise credibility.

Table 2.2: Categories for beliefs about knowledge in chemistry

Epistemology Dimension	Naive Viewpoint
Simple	Most questions have one right answer
Certain	Knowledge is unchanging
Authority	Teachers and scientists know the answers
Truth	Truth is attainable

With regard to epistemological beliefs and feedback in this study, we anticipate that students having a high belief in the simple/certain dimension of chemistry knowledge will be less likely to read all of the feedback, or may prefer simple feedback and be "put off" by more lengthy feedback. We anticipate that this may be revealed in differential performance between simple and complex problems, as complex problems tend to require complex explanations in the way of feedback. We also anticipate that students with a strong belief in the attainability of truth and/or strong belief in authority may take feedback as “the” truth, choosing to memorize the explanation rather than accommodating the new information.

There is another potential effect involving epistemological beliefs and that is in the influence of feedback on beliefs. Prior research shows that one’s epistemological beliefs can and do change, albeit over long periods of time (Schommer, 1993, 1998). These changes are believed to be induced by academic challenges and academic settings. It is possible, though improbable, that feedback may influence a student’s epistemological beliefs if given enough time and treatments. In addition, diagnostic feedback, in its explanation of why an alternate answer is better than the response chosen, may assist a change in a student’s simple/certain beliefs more than explanation alone (informational feedback). Lastly, if diagnostic feedback can help the student to believe that she can weigh the facts and make judgments just as well as the “experts” then that student’s belief in authority may be shifted.

This chapter discussed feedback and its relationship to performance and motivation and proposed a potential relationship with epistemological beliefs. With these variables in mind, our overarching research question is:

> Does informational or diagnostic feedback have the greater positive influence on performance, achievement goal orientation, and epistemological beliefs toward chemistry?

With a supplemental research question:

> What is the relationship among the variables of feedback type (specificity), feedback count, performance, motivation, and epistemological beliefs in chemistry?

Chapter 3: Methodology

This research involved an experimental field study comparing two feedback strategies, based on specificity of information in the feedback. The research occurred during the second semester of the general chemistry sequence at a large southern research university. The method employed for the study was largely quantitative in nature—a modified pre-test, post-test control group experimental design. The focus was on the comparison of two types of cognitive feedback—informational feedback and diagnostic feedback—on performance, motivation, and epistemological beliefs. The feedback was delivered via online quizzes associated with the lecture course. It is important to note that the two feedback types differed only when the student missed a quiz question. Thus, a differential treatment was given only for missed questions.

This chapter first presents information about a brief pilot study that was conducted, then describes the various aspects of the research study including the sample, measures, and data analysis.

Pilot Study

Longitudinal studies, even spanning only four months, require careful planning. Online education, including online quizzes used in this study, also requires time for “working out the bugs” and investigating potential pitfalls. Therefore, a pilot study was conducted during the Spring 2003 semester.

The pilot study involved students reading one-by-one individual quiz questions from the computer, responding to each multiple-choice question, and reading the resulting feedback from the computer screen before continuing to the next question. The primary purposes of the pilot study included the investigation of the value of categorizing questions according to an adapted Bloom’s taxonomy (Bloom, 1968), and testing the viability of quizzing online. Another concern involved the web-based interface and potential problems the students might encounter with it while completing electronic assignments, so two pilot groups were observed in a computer laboratory environment and trouble areas noted so that they could be minimized in the full study.

Outcomes

After the pilot study, discussion with student participants and colleague observers influenced this study’s version of the e-quiz format and interface. Watching participant behavior during the pilot also raised another question, which would lead to hypotheses regarding epistemology.

In our preliminary study, students did indeed employ different strategies in completing the various problems and they performed less well on the higher-classified problems, providing evidence that they viewed them as being more challenging (in line with prior researchers' assumptions along this line, see for example Bloom, 1968). Thus, this study continues the use of categorizing questions according to difficulty, based on Bloom's taxonomy (1968). Also, the majority of the students proved able to successfully complete online assignments without trouble. However, a few students found "minor" problems like typographical errors or technological inconsistencies as being paralyzing, which moved the researcher to both pay extra attention to errors and to investigate a parameter that might predict such debilitating response in the online environment. This investigation led us to the notion of epistemological beliefs, which opened other opportunities and questions.

Because of the pilot study observations, we chose to perform the full study with online quizzes, using immediate feedback, categorizing questions according to difficulty to determine whether one feedback type would assist in a particular level of question more than another feedback type, and we chose to include motivation and epistemological beliefs as measures, with the hopes of determining whether these measures influence the acceptance of feedback. It was during the design of the feedback that we wondered whether feedback could, in turn, influence motivation and epistemological beliefs.

Sample

The full-study sample consisted of students enrolled in one section of the second semester of general chemistry. Students in the course in spring 2004 participated in the experiment. It should be noted that the professor for this section has won several teaching awards and is known as an excellent instructor.

Enrollment in this 8 A.M. section exceeded 200 students who met for lectures three times a week for fifty (50) minutes each session. A voluntary attendance “group office hours” section was also held weekly on Tuesday evenings, after the online quizzes (“e-quizzes”) closed. Two teaching assistants and the professor were available at these office hours.

Since the e-quizzes were a part of the course, all students participated; however, only the 95 adult students who gave consent and completed all surveys were included in the study analyses. The student sample thus consisted of 41 males and 54 females, with age distributions as shown in Table 3.1. Only 41 of the 95 students in the sample were classified as freshman (Table 3.2). Students under 18 years of age as of

September 1, 2003, were removed from the sample, according to our IRB contract for this study. Of the sample set, 48 students received informational feedback and 47 received diagnostic feedback.

Table 3.1: Age Distribution of Participants Spring 2004

Age (as of Sept. 1, 2003)*	Number, n
18-20 years	64
20-22	19
22-25	8
>25	4

*Self-reported age by survey.

Table 3.2: Year in College of Participants Spring 2004

Year in College	Number, n
New freshman (first semester)	31
Transfer freshman	3
Other freshman	7
Sophomore	27
Junior	11
Senior	14
5th year senior	2

Self-reported year in school as of September 1, 2003.

Independent Variables

Two independent variables were assigned in this study: feedback type and treatment count; that is, the specificity of feedback and the number of feedback events, respectively.

Feedback type

Student participants were randomly assigned to one of the two feedback types for the duration of the study: informational feedback (IF) or diagnostic feedback (DF). The students took online quizzes and received feedback immediately after responding to each question. The information given in the feedback depended on which group the student was assigned, either informational feedback or diagnostic feedback.

The informational feedback recipients received an “Incorrect” or “Correct” notation followed by an explanation of the problem. The diagnostic feedback recipients received feedback keyed to their response. In writing the multiple-choice question, the distracter was specifically selected to diagnose an error and the diagnostic feedback was designed to confront that error and correct it (a “confront and correct” implementation of feedback). Thus, in the case of a correct response, the diagnostic feedback exactly matched that of the informational feedback (a notation of “Correct” plus an explanation of the problem) because no detectable error existed to correct. However, in the case of an incorrect response, the diagnostic feedback included a notation of “Incorrect”, ***plus*** a statement of the error, followed by the explanation of the problem. Thus, the principal difference between the informational and diagnostic feedback types was the corrective statement of what the student did wrong. It is very important to notice that the two feedback types in this study differed only for incorrect responses. A sample problem followed by examples of each feedback type is shown in Figure 3.1. (More examples are available in Appendix A.3.)

Figure 3.1 Sample Question and Feedback

Sample Question and Feedback

A piece of shiny metal has a mass of 64.37 g and a volume of 8.2 cm^3. Calculate its density.

A. 7.850 g/cm^3

B. 527.83 g/cm^3

Informative Feedback Examples

If the student chooses option A, then he sees:

That is correct! Good job. Density is the ratio of mass per volume. It represents how much weight (mass) is in a certain volume. In order to find density, you must take the mass and divide it by the volume.

If the student chooses option B, then he sees:

That is incorrect. Density is the ratio of mass per volume. It represents how much weight (mass) is in a certain volume. In order to find density, you must take the mass and divide it by the volume.

Diagnostic Feedback Examples

If the student chooses option A, then he sees:

That is correct! Good job. Density is the ratio of mass per volume. It represents how much weight (mass) is in a certain volume. In order to find density, you must take the mass and divide it by the volume.

If the student chooses option B, then he sees:

That is incorrect. It looks like you incorrectly multiplied the two numbers to get the answer. But density is the ratio of mass per volume. It represents how much weight (mass) is in a certain volume. Mass is 64.37 g and volume is 8.2 cm^3. In order to find density, you must take the mass and divide it by the volume. Divide mass by volume to get density.

Treatment Count, TC

Because feedback differed between the two types only for questions missed, it may be said that differential treatment occurred only for missed questions. The Treatment Count (TC) was established mathematically as the running sum of the number of e-quiz questions **missed** across all e-quizzes. Conceptually, it represents the number of treatments each student encountered. The nature of diagnostic feedback used in this study mandated this approach: a diagnosis was made only when the student's response was incorrect. So, when a student in the diagnostic group answered a question correctly, that student received informational feedback. Students in this study who were assigned to the diagnostic treatment group only received differential treatment when they answered a question incorrectly. For purposes of comparison, the TC was calculated for the informational group as well.

Dependent Variables

Three sets of dependent variables were measured in this study, with the purpose of investigating which, if any, related to feedback type or treatment count as well as providing data for other, unanticipated correlations between the measures themselves. The following paragraphs discuss the dependent variables of performance, motivation orientation, and epistemological beliefs toward chemistry and their respective instruments.

Performance Measures

Several performance measures were used, depending on the nature of the research question. The final course score, in numerical form as the ratio of points received per total points available, was obtained from the course professor minus any contribution from the e-quizzes. The course score was an average score, reported as a ratio and based only on examination scores.

Other performance scores came from the e-quizzes themselves. It is important to note that the e-quizzes were voluntary. The students had a possibility of gaining "extra points" if they passed either one of the two e-quizzes offered each week. The students "passed" an e-quiz if they got at least two-thirds of the questions correct. This point is reiterated because scattered participation was observed and in order to compensate mathematically for this factor as well as the fact that different e-quizzes had different numbers of questions (9-12 questions each), simple averages as scores were not used in the analysis. Instead, each score was calculated as the ratio between number of questions correct to the number of questions attempted. Calculating the overall score in this way skews the value away from a “percentage” performance in the traditional sense, but allows inclusion of all participants, whether they took only one e-quiz or all of them, only one e-quiz each week or both, and allows each question to be counted individually and with equal weight as opposed to averaging across e-quizzes having different numbers of questions. The following three e-quiz scores were calculated in this manner.

The <u>overall e-quiz score</u> was calculated as the number of e-quiz questions correct divided by the number of e-quiz questions attempted during the entire semester. In a similar manner, the analysis of performance over time used <u>e-quiz unit scores</u>. These scores were calculated by dividing the number of questions correct by the number of quiz questions attempted during that unit. A unit is the work the student did from the last in-class examination to the next in-class examination. For example, before the first examination, four e-quizzes were given across two weekends. Each student’s e-quiz score for unit 1 (EQS_U1) was therefore calculated as the number of quiz questions correct divided by the number of questions attempted on the four (4) e-quizzes preceding the first examination.

Performance by difficulty level was also calculated. In keeping with Bloom's taxonomy, (1968), every e-quiz question was categorized as either rote (R), comprehension (C), or application/analysis (A). The e-quiz score for each level of difficulty was calculated per student by dividing the number of questions correct in that category by the number attempted. Thus, the e-quiz score for rote-level questions (EQS_R) was calculated as the ratio of the number of rote-level questions correctly answered per rote-level questions attempted.

Qualitative-type analysis was conducted across individual questions. Performance on a per-question level was binary: (1) correct response or (0) incorrect response.

Motivation Instrument

Motivation was determined by an instrument employed recently by Milner-Bolotin (2001), which is based on the Midgley, et al, Achievement Goal Orientation Questionnaire (1998). The Milner-Bolotin instrument is a paper-and-pencil survey consisting of 15 Likert-style statements. Previously validated in the college-level classroom for physical science students at The University of Texas at Austin (alpha=0.83), the changes from the original version of the instrument included changing terms "physical science" to "chemistry" and "project work" to "assignments" (The survey is located in Appendix A.1). The alpha for this study's sample was 0.79 and for an independent non-science sample at this same institution, alpha was 0.78. The mastery, performance-approach, and performance-avoid orientations are viewed in the literature as being independent of one another; so a principal components analysis was conducted to determine factor assignments. Assignment of questions to goal orientations is located in Appendix A.2.

The survey captured four factors for the motivation measure: (1) mastery goal orientation (MGO), (2) performance-approach goal orientation (SHOW), and (3 & 4) two sub-categories within performance-avoid orientation, namely, avoid embarrassment goal orientation factors 1 and 2 (AV1, AV2). Achievement goal theory identifies a student's "purposes for engaging in academic work" (p. 127 Midgley, et al, 1998). The mastery orientation (MGO) factor determined the extent to which the student has goals to develop skills or knowledge or tends to find satisfaction in learning. The performance-approach (SHOW) factor determined the extent to which the student desired to outperform others or demonstrate proficiency. The first performance-avoid factor (AV1) determined the extent to which the student desired to avoid looking incompetent. The second performance-avoid factor (AV2) differed slightly in that it determined the extent to which the student desired that others not think him or her ignorant or less knowledgeable. A student may have strengths in more than one area simultaneously.

Responses to questions loading on the same factor were averaged per student to acquire the student's score on that subscale. Thus, the mastery goal orientation factor score is the average of the responses to the four survey questions (1=strongly disagree, 5=strongly agree):

> I will do the chemistry assignments because I am interested in them.
>
> I like working on more difficult quiz questions that I'll learn more from, even if I make a lot of mistakes.
>
> An important reason I participate in the chemistry assignments is because I want a deeper understanding of chemistry.
>
> I like assignments best when they really make me think.

Epistemological Beliefs Instrument

The epistemological beliefs instrument used in this study was based on Hofer's study on subject-specific epistemological perspectives, which included science (Hofer, 1997). Hofer's instrument is a "3rd generation" survey based on Qian and Alvermann's instrument (1995), which originated from the very popular epistemology instrument by Schommer-Aikins (1990, 1998). The statements in Hofer's instrument were specifically written to measure epistemological beliefs toward science. It is a paper-and-pencil instrument focusing on constructs that overlap across epistemological models and form the “core of an individual’s personal epistemology” (p.91, Hofer, 1997). Just as Qian & Alvermann took select questions from Schommer’s instrument, thereby fine-tuning the survey, and just as Hofer took select questions from Qian & Alvermann’s instrument, a selection of questions from the Hofer instrument was used for this study. Selection of questions was based on Hofer’s findings that a group of her questions did not correlate with performance and that performance was most strongly correlated with a particular set of questions stemming from the simple/certain dimension.

The survey used in this study is in Appendix A.1. During completion of the instrument, students were asked to keep chemistry and learning chemistry in mind as they answered the Likert-style statements numbering from 1=”strongly disagree” to 5=”strongly agree”. The alpha for this study’s sample was 0.75. The alpha for an independent sample at this same institution was 0.74.

The survey captured three factors for the subject-specific epistemological beliefs measure, based on Hofer’s (1997) factor structure: (1) the simple/certain (CERT) dimension, (2) the belief in authority (AUTH) dimension, and (3) the attainability of truth (TRU) dimension. Epistemological beliefs, as they apply to this study, attempt to measure a student's cognitive filters in receiving, interpreting, and believing information. Each dimension ranges along a continuum, from naïve to mature beliefs about knowledge . A naïve perspective in the simple/certain dimension would be measured by a student's belief that truth is unchanging, that there is only one answer to any question, and that principles in a field are unchanging or static. The authority dimension measures the extent to which a student believes that persons in authority have "all the answers in a given field". These students would tend to believe textbooks rather than personal experiences when learning a subject. The attainability of truth dimension measures the extent to which the student believes that truth can be found by experts or scholars.

These measures may or may not be independent of each other and students can hold mature beliefs in one dimension and naïve beliefs in another. A score was calculated for each student on each dimension, based on averaging Likert-scale responses. Questions were assigned to dimensions based on research conducted by Hofer (1997). The survey and factor assignments for this study are in Appendix A.2

Procedures

The study was conducted during the academic year 2003-2004. The study took place during the second semester of the general chemistry sequence. Content included the topics of gases, liquids and solids, solubility, thermodynamics, chemical kinetics, equilibrium, acid/base equilibrium, buffers, and electrochemistry. The class met at 8am on Mondays, Wednesdays, and Fridays for 50-minutes each session.

Surveys

The initial survey was conducted during the first two weeks of the Spring 2004 semester. The final survey was conducted during the last week of classes in Spring 2004. The packet consisted of two surveys: the first to measure motivation goal orientation and the second to estimate epistemological beliefs, and a scantron sheet on which the students submitted their responses. The final survey packet also included multiple-choice demographic questions. Any student who missed the class when the surveys were administered was invited to take them at set times spanning several days, up to a week after the class-administered survey. These surveys are located in Appendix A.1.

Consent

Consent to participate was collected from the participants at the end of the spring semester. The students were informed that the study was an investigation of feedback on performance and that the researcher was interested in finding better ways to teach chemistry using online quizzes and feedback. The students were reassured that analysis would take place after their grades were reported to the registrar and that their names would be removed from all data before analysis. The relevant IRB consent form appears in Appendix A.4.

E-Quizzes and Feedback Types

Each student was randomly assigned a “code” by the computer software that designated the feedback type that student would be given for the duration of the study. To enter the feedback information into the software for each question, the researcher used several text-entry box triads. A triad consisted of three text-entry boxes: one for the answer choice, one for the informational feedback, and one for the diagnostic feedback. One triad existed for each multiple-choice option in the question—usually four or five, consisting of the correct answer and three or four distracters. The informational feedback text was copied and pasted from one answer choice triad to all of the others such that all students receiving informational feedback were presented with the same text (except for altering “Correct” to “Incorrect” as appropriate). On the other hand, diagnostic feedback changed depending on the answer response chosen by the student. This allowed for different choices to result in different feedback presented. However, it is noteworthy that sometimes students in the diagnostic feedback group received informational feedback, which happened when the student answered the question correctly. Since this study employs the “confront and correct” interpretation of diagnostic feedback, the student in the diagnostic feedback group who answered the question correctly had nothing (detectable) to diagnose or correct, so she received informational feedback.

The online quizzes (e-quizzes) were posted once each week, normally on Fridays, and were due the following Tuesday before the weekly recitation section. The e-quiz content covered that week’s lecture instruction, which usually included multiple topics of general chemistry. A typical e-quiz consisted of six to twelve questions of increasing difficulty, starting with rote-level, followed by equal numbers of comprehension and application- or analysis-level questions. Students were required to complete each quiz in one sitting (all at once) at a computer of their choice. Computer-recorded start-end times for students show that students typically finished each e-quiz in 30-45 minutes, with some students taking an hour. The students were asked to treat these assignments as quizzes by not consulting their friends or textbooks; however, this condition was not policed.

Two e-quizzes were available each week. In order to encourage participation in the e-quizzes but not to force it upon the students, the professor told the students that the results of the e-quizzes would not count for a set percentage of the grade, but that they could potentially add up to 2% (0.02 added to the ratio score) for students on the edge between letter grades (e.g. a student receiving 0.88 final course score and who passed an e-quiz each week could receive an A in the course). The students needed to correctly answer two-thirds of the problems on one of the two e-quizzes each week for the work to be counted as a “pass”.

To begin an e-quiz, the student accessed to the course website from any computer connected to the Internet, clicked the “E-quizzes” link, read the brief introduction identifying the number of questions and content topics, and then clicked a link to begin the e-quiz. After logging in with a unique ID and password, the student saw the first question and its multiple-choice responses, each with a radio button (a click-able circle) next to it. After the student selected a response by clicking on a button (circle), text-based feedback appeared beneath the multiple-choice responses and a “next” button appeared as well, to guide the student to the next question. The feedback was therefore provided immediately after the student’s response to each question and the student controlled when the next question would be viewed.

It is important to note that the students had to take the e-quiz in a linear manner, which means that they had to answer question 1 first, then question 2, etc. They could not go back to change their answer because the response had already been recorded and feedback given. However, both the question and answer choices, with the student’s response shown in bold-faced type, were available during presentation of each problem’s feedback.

Data Analysis

Data Analyses consisted of two general types: quantitative and in-depth qualitative. The quantitative methods give a statistical comparison of the two feedback types and treatment count (TC) on the dependent variables. Comparison of means and follow-up t-tests investigated the difference between means for each feedback type. Repeated measures were used to investigate pre-post survey changes and differences between feedback types. Correlation was used to determine what relationships existed between the variables and thus which relationships warranted further study. Plots of the data and nonlinear regression analyses were used to investigate the relationships detected with correlation.

The in-depth qualitative-type analysis allowed the researcher to get a per-question view of the effect of feedback on individual students at particular snapshots in time. This approach to the data gave another vantage point for studying feedback and subsequent performance on same and similar multiple-choice questions and mistakes that individual students made before and after receiving the feedback.

Regression analysis was performed to determine if initial motivation or epistemological beliefs affected the performance measures. To analyze feedback specificity and performance, three performance measures were used and statistical tests involved either t-tests or repeated measures, as applicable. To analyze feedback specificity and motivation, pre-post survey scores were used so repeated measures were employed. Analysis of feedback specificity and epistemological beliefs closely resembled the motivation analysis; pre-post survey scores and repeated measures were used. To investigate whether the measures performance, motivation, or epistemological beliefs varied with treatment count (TC), a correlation matrix was calculated first for the entire sample, then separately for each feedback type. Scatter plots and other graphs aided the interpretation, as well. To investigate whether students reproduced their mistakes and to determine whether information given in feedback was retained or used by students in transfer problems, a more qualitative approach was employed as success rates were analyzed on a per-question level and compared across the two feedback types.

The results to these inquiries are presented in the next chapter.

Chapter 4: Results

Two main goals drive this study: to explore electronic quizzing for helping students learn chemistry, and to determine whether diagnostic feedback meets that task better than informational feedback. In working toward these goals, two overarching objectives were formulated: (1) to compare two feedback types and (2) to determine more information about the relationships among the variables of feedback type, feedback count, performance, motivation, and epistemological beliefs toward knowledge in chemistry. Thus, the impetus for this research includes both research and application motives.

This chapter presents results relevant to the many questions stemming from the objectives above, including quantitative and qualitative data. The variables are listed and defined in Table 4.1. The data are categorized in Table 4.2. The quantitative research questions, measures, and statistical tests are summarized in Table 4.3. The questions are organized by the variables “feedback type”, “feedback count” (number of exposures to the corrective feedback), “performance”, “motivation”, and “epistemological beliefs”. Note that the statistical test chosen depends on the nature of both the variables included and the nature of the question posed. The in-depth analyses, based on more qualitative-style methods, are summarized in Table 4.4.

Because past research has shown performance to be related to both motivation and epistemology, the first questions will examine whether there is a correlation between these measures. If there is, then these variables will need to be used as covariates in the subsequent analyses of performance with feedback type.

The next research questions look at whether performance, motivation and epistemological beliefs in chemistry vary with feedback type and whether there is a significant difference between the two feedback types on any of these measures. Also, since this is a longitudinal-type study, it provides the rare opportunity to observe the effects of number of feedback exposures (treatment count, TC) on performance, motivation, and epistemological beliefs, as presented in the last of the quantitative-based research questions.

Finally, the results of the quantitative analyses generated questions about whether and how students were using the feedback provided and whether a difference would be observed between the two feedback types for paired quiz questions; the quasi-qualitative research questions were formulated to explore these areas.

Table 4.1: List and Definitions of Variables Used in Study

Abbrev.	Variable	Definition
IF	Feedback type	Informational feedback type (what, why, how)
DF	Feedback type	Diagnostic feedback type (why not, what, why, how)
TC	Feedback count	Number of “treatments” received; also is number of questions missed
U1, U2, U3, U4	Units	Groupings of content matching examinations. U1 includes e-quizzes given up until examination 1, U2 for e-quizzes up to examination 2, etc.
EQS	Performance	E-Quiz score. Ratio of number of quiz questions answered correctly to number of quiz questions attempted, for a given student
CS	Performance	Course score. Simple average of four examination scores for a given student
RCA	Performance	R=rote, C=comprehension, A= application & analysis level of question difficulty
MGO	Motivation	Mastery goal orientation score resulting from an average across survey questions.
SHOW	Motivation	Performance-approach goal orientation. Score is an average across survey questions.
Av1 & Av2	Motivation	Performance-avoid goal orientation subscales. Score is an average across survey questions for each subscale.
CERT	Epistemology	Certain/Simple dimension. Score is an average across survey questions.
AUTH	Epistemology	Authority dimension. Score is an average across survey questions.
TRU	Epistemology	Attainability of truth dimension. Score is an average across survey questions.

Table 4.2: Overview of Research Measures and Components of Study

	Feedback Count, TC	Performance			Motivation	Epistemology
		EQS	CS	RCA		
Instrument	E-quizzes	E-quizzes	Class examinations	E-quizzes	Milner-Bolotin Motivation Survey	Hofer Epistemology Survey
Type of Data Collected	Quantitative; Count of number of questions missed	Quantitative; Ratio correct responses to number of questions attempted	Quantitative; Average of examination scores in classroom	Quantitative; Ratio of correct responses to number of questions attempted, per level of difficulty category	Quantitative; 15 Likert scale responses	Quantitative; 14 Likert scale responses
Type of Scores Produced	One overall count	One overall score and four per-unit scores	One overall numerical (%) score	Three scores, one for each level of difficulty: R, Rote, C, Comprehension, and A, Application/Analysis	Four factors: MGO (mastery), SHOW (performance-approach), and Av1 & Av2 (performance avoid)	Three factors: Cert (belief that knowledge is simple and/or certain) Auth (belief that knowledge comes from authority), Tru (belief that truth is attainable)

TC: Treatment Count; EQS: E-Quiz Score; CS: Course Score; RCA: level of question (R, rote; C, comprehension; A, application or analysis); MGO: Mastery Goal Orientation; SHOW: Performance-approach Goal Orientation; Av1 & Av2: Performance Avoid orientation factors; Cert: Certain dimension; Auth: Authority dimension; Tru: Attainability of Truth dimension

Table 4.3: Quantitative Analyses

Question	Instruments or Data Collection Methods					Variables	Statistical Analysis
	EQS	CS	M-B	HOF	TC		
Initial Motivation and Performance							
Q0a: Does initial motivation predict performance?	Overall ratio score	Overall ratio score	MGO, SHOW, Av1, Av2			IV: MGO, SHOW, Av1, Av2 DV: CS, EQS	Correlation
Initial Epistemology and Performance							
Q0b: Do initial epistemological beliefs of chemistry predict performance?	Overall	Overall numerical score		Cert, Auth, Tru		IV: Cert, Auth, Tru DV: CS, EQS	Correlation
Feedback Type and Performance							
Q1a: Do students who receive DF achieve a higher course score in chemistry than those who receive IF?		Overall numerical score				IV: FT DV: CS	Means, std. error t-tests
Q1b: Do students who receive DF exhibit better e-quiz performance over time than those who receive IF?	Unit 1, Unit 2, Unit 3, Unit 4					BET: DF, IF WIN: EQS	Repeated measures
Q1c: Does DF differ from IF on performance of questions of different levels of difficulty?	R C A					IV: DF, IF DV: EQS on R, C, A categorized problems	Means, std. error t-tests
Feedback Type and Motivation							
Q2: Do DF recipients become more motivated than their IF counterparts?			MGO, SHOW, AV1, AV2			BET: DF, IF WIN*: MGO, SHOW, AV1, AV2	Repeated measures*

Question	Instruments or Data Collection Methods					Variables	Statistical Analysis
	EQS	CS	M-B	HOF	TC		
Feedback Type and Epistemological Beliefs							
Q3: Do DF recipients change their epistemological beliefs about the knowledge of chemistry more than their IF counterparts?				Cert		BET: DF, IF WIN*: Cert, Auth, Tru	Repeated measures*
Feedback "Treatment" Count (TC)							
Q4a: Do performance, motivation, or epistemological beliefs of chemistry change with TC?	Overall	Overall numerical score	MGO, SHOW, Av1, Av2	Cert, Auth, Tru	Sum	All	Correlation
Q4b: What is the relationship among DF, TC, motivation, and epistemological beliefs of chemistry?	Overall	Overall numerical score	MGO, SHOW, Av1, Av2	Cert, Auth, Tru	Sum	All, but data only for DF	Correlation
Q4c: What is the relationship among IF, TC, motivation, and epistemological beliefs of chemistry?	Overall	Overall numerical score	MGO, SHOW, Av1, Av2	Cert, Auth, Tru	Sum	All, but data only for IF	Correlation

*Individual repeated measures analysis was conducted for each within-subjects variable.
DF: Diagnostic Feedback; IF: Informational Feedback; FT: Feedback type (dummy variable); BET: Between-subjects variable; WIN: Within subjects variable; EQS: E-Quiz Score; CS: Course Score; M-B: Milner-Bolotin motivation orientation instrument; HOF: Hofer Epistemological beliefs instrument; TC: Treatment Count; IV: Independent Variable(s); DV: Dependent Variable(s); MGO: Mastery Goal Orientation; SHOW: Performance-approach Goal Orientation; Av1 & Av2: Performance-Avoid orientations; Cert: Certain/Simple dimension; Auth: Authority dimension; Tru: Attainability of Truth dimension

Qualitative Analysis Table

Table 4.4: Quasi-qualitative Analysis

Question	Measure or Data Analyzed	Method of Comparison
Q5: Do students who receive DF outperform their IF classmates on a same or similar question posed (1) soon after initial exposure and correction and (2) later, on an examination?	Number of students correctly answering each question in each of three case studies, per group (DF, IF).	Construction of tabular data and invention of coding technique to identify patterns of successful and unsuccessful responses to e-quiz questions in each case study.
Q6: Do students who receive feedback reproduce their mistakes? Are DF recipients less likely to make the same mistake again than their IF counterparts?	Each student's individual response (content of response) to each question in each of three case studies.	Analysis of each student's response to first question, followed by individual response to second and third viewings of same question. Search for any trends in the data, on a per-individual basis. Analysis of possible reasoning behind a student's choice of distracter and its relationship to the problem's solution.

Quantitative Analysis

Initial Motivation and Epistemological Beliefs

Q0a: Does initial motivation predict performance?

Q0b: Do initial epistemological beliefs in chemistry predict performance?

Initial motivation was measured by survey during the second week of class, as described in Chapter 3. Achievement goal orientation, as used in this study via Milner-Bolotin's instrument, includes four subscales: MGO, SHOW, Av1, Av2 (defined in Table 4.1). Performance includes five variables: course score, e-quiz performance, and the three R, C, A scores (defined in Table 4.1).

Students' epistemological beliefs in chemistry were captured by Hofer's (1997) survey that was specially designed for students in the sciences. We include survey questions from three subscales: beliefs that chemistry knowledge is certain (unchanging) and/or simple, belief in authority as the primary source of knowledge, and belief that attainability of truth is possible in chemistry. Student performance included five variables: course score, overall e-quiz score, and the three subscales of R, C, A, performance.

The correlation results show that one subscale, mTRU significantly changes with the performance variable overall e-quiz performance. Since the correlation is less than 0.5, the subscale will not need to be employed as a covariate. No other pre-treatment survey subscale correlates with either performance measure, so no other will need to be used as a covariate. This finding is similar to a longitudinal feedback study by Harackiewicz, et al., (2000), which also reports that mastery goals measured in the beginning of the semester did not predict performance.

As an aside, the significant and somewhat strong correlation between e-quiz score and course score reflects the effort that went into selecting questions for the e-quiz that matched the topics in the classroom and the instructor's style, manifested in questions on the course exams. It would seem that the effort paid off; a high score on the e-quizzes corresponds with a high course score—something that will be discussed further later in this report.

Table 4.5: Pearson Correlations for Initial Motivation and Performance

		course score	EQS_TOT	MMGO	MSHOW	MAV1	MAV2
course score	Pearson Corr.	1.000					
	Sig. (2-tailed)	.					
	N	95					
EQS_TOT	Pearson Corr.	.440	1.000				
	Sig. (2-tailed)	.000	.				
	N	87	87				
MMGO	Pearson Corr.	.061	.050	1.000			
	Sig. (2-tailed)	.560	.649	.			
	N	95	87	95			
MSHOW	Pearson Corr.	.004	-.167	.008	1.000		
	Sig. (2-tailed)	.968	.121	.941	.		
	N	95	87	95	95		
MAV1	Pearson Corr.	.044	.125	.087	.455	1.000	
	Sig. (2-tailed)	.674	.248	.401	.000	.	
	N	95	87	95	95	95	
MAV2	Pearson Corr.	.091	.056	-.101	.217	.495	1.000
	Sig. (2-tailed)	.383	.605	.332	.035	.000	.
	N	95	87	95	95	95	95

Table 4.6: Pearson Correlations for Initial Epistemological Beliefs in Chemistry and Performance

		course score	EQS_TOT	MCERT	MAUTH	MTRU
course score	Pearson Corr.	1.000				
	Sig. (2-tailed)	.				
	N	95				
EQS_TOT	Pearson Corr.	.440	1.000			
	Sig. (2-tailed)	.000	.			
	N	87	87			
MCERT	Pearson Corr.	-.147	-.042	1.000		
	Sig. (2-tailed)	.156	.696	.		
	N	95	87	95		
MAUTH	Pearson Corr.	.004	.008	.454	1.000	
	Sig. (2-tailed)	.968	.943	.000	.	
	N	95	87	95	95	
MTRU	Pearson Corr.	-.156	-.244	.315	.127	1.000
	Sig. (2-tailed)	.130	.023	.002	.219	.
	N	95	87	95	95	95

Feedback Specificity and Performance

The purpose of this first set of research questions is to analyze whether diagnostic feedback is “better” than informational feedback, with “better” demonstrated by higher performance scores.

Q1a: Do students who receive diagnostic feedback achieve a higher course score in chemistry than those who receive informational feedback?

The purpose of this question is to examine the difference between feedback types for overall course score. A comparison of means and standard errors showed no significant difference between the diagnostic feedback group and the informational feedback group in course score (see Table 4.5)(t=-0.472, *n.s.*).

Table 4.7: Descriptive Statistics for Overall Course Score

Type	Informational	Diagnostic
n	48	47
Mean Course Score	0.81907	0.82970
s.d.	0.11933	0.09926
std. error	0.01722	0.01448

Thus, it was found that diagnostic feedback does not lead to better overall performance in chemistry beyond that obtained through informational feedback, as measured with overall course score. However, it is possible that students may be able to maintain adequate performance in a course regardless of instruction given, if they study outside of class or have other sources of input. To see the effect of feedback on subsequent performance, perhaps it is better to look only at e-quiz performance rather than overall course score.

Q1b. Do diagnostic feedback recipients exhibit better e-quiz performance over time than those who receive informational feedback?

For the second performance measurement—e-quiz performance over time—repeated measures were used to determine if there was a time-related change in performance. The hypothesis is that perhaps the information in the feedback will trigger some knowledge or perhaps scaffold the student's learning through explanation and modeling of strategies such that future improvement will be observed, as suggested by other researchers (Annett, 1969; Kulhavy, 1977).

Repeated measures were used to determine whether performance changed over time and whether that change was different for each feedback type. The variable PERFORM was employed to represent the e-quiz scores for individual units. The use of units as time points allowed for the inclusion of students in the study who missed an occasional e-quiz. In such a case, the unit average contains only information for e-quizzes in which the students participated. The units corresponded with the examination schedule; unit 1 included content covered until the first examination, unit 2 content up to the second examination, and so on. A student's score for a given unit was computed as the ratio of the count of the number of questions answered correctly divided the number of questions attempted during that time period. Unit 1 included up to 4 e-quizzes, Unit 2 included up to 4 e-quizzes, Unit 3 included up to 5 e-quizzes (one week offered only one version), and Unit 4 included up to 4 e-quizzes. Recall that each e-quiz consisted of 6 to 12 questions of varied difficulty.

The results of these analyses (Tables 4.8-10) show change in e-quiz performance across the four time points (units) for both groups in that both groups missed increasing numbers of e-quiz questions as the semester progressed. The diagnostic feedback group missed fewer questions in each unit than the informational feedback group, but the difference is not significant.

Table 4.8: E-quiz Scores[a,b] over Units (time) and Feedback Types

		Informational	Diagnostic
	n	31	24
EQS_U1			
	Mean	0.715	0.743
	Std. Error	0.031	0.035
EQS_U2			
	Mean	0.613	0.635
	Std. Error	0.028	0.032
EQS_U3			
	Mean	0.679	0.696
	s.d	0.021	0.024
EQS_U4			
	Mean	0.561	0.595
	Std. Error	0.035	0.040

[a]These means represent the 55 cases used in the repeated measures study.
[b] Evaluated at covariates appeared in the model: MCERT = 2.51, MAUTH = 3.06.

Table 4.9: Tests of Within-Subjects Effects for E-Quiz Performance over Units

Source	SS	df	MS	F	Sig.	Eta Squared
PERFORM	.726	2.395	.303	15.219	.000	.223
PERFORM * GROUP	1.905E-03	2.395	7.953E-04	.040	.977	.001
Error(PERFORM)	2.527	126.955	1.991E-02			

Table 4.10: Tests of Between-Subjects Effects for E-Quiz Performance over Units

Source	SS	df	MS	F	Sig.	Eta Squared
Intercept	92.747	1	92.747	1596.036	.000	.968
GROUP	3.539E-02	1	3.539E-02	.609	.439	.011
Error	3.080	53	5.811E-02			

The significant main effect for PERFORM in Table 4.9 indicates that e-quiz performance changed significantly between the units. That is to say that the e-quiz performance changed over time for the sample as a whole in this analysis. The lack of significance for the PERFORM*GROUP interaction in the within-subjects effects indicates that change in performance over time (slope of plot of e-quiz score by unit) was the same for both feedback types (groups).

The non-significant main effect for GROUP in the between-subjects analysis (Table 4.10) indicates that there was no significant difference between the e-quiz performances of the two feedback types. Thus, the results from the repeated measures analysis reveals that increased feedback specificity did not have a differential effect on e-quiz performance over time.

Q1c. Does performance differ with feedback type and varied levels of difficulty of questions?

Some researchers have asserted that more specific feedback may be more beneficial for complex tasks (Jacoby, 1984; Moreno, 2004) than easy tasks. In order to analyze this in the chemistry classroom, we classified the e-quiz questions according to Bloom’s taxonomy (Bloom, 1968), ranging from rote-memory (R) to comprehension (C), to application/analysis (A). This made it possible to examine gradations of question difficulty as they relate to performance for each feedback type. A comparison of means was again employed, this time focusing on performance per level of question difficulty. Performance per level was computed as the ratio of the number of questions in that category successfully answered to the number of questions attempted in that category.

Means, standard deviations, standard errors and t-test results are presented in Table 4.11. Since some students in the sample did not participate in any e-quizzes, the sample size for each is reduced compared to the entire study sample.

The sign of the t-test gives an idea of which feedback type "outperformed" the other, though it does not denote significance. In the case of a positive t-value, the informational feedback mean is greater than the diagnostic mean and so the students receiving informational feedback outperformed their diagnostic feedback counterparts. This occurred for the rote-level questions, though it is not even nearly significant. The opposite occurred for the higher-level questions, but again the differences are not significant.

Therefore, the two groups performed similarly on e-quiz questions of rote level, comprehension level, and application/analysis level questions, though their performance was significantly poorer for higher-level questions than for rote-level questions when cross-level t-tests were conducted.

Table 4.11: Comparison of Means Across Levels of E-quiz Question Difficulty

		Informational	Diagnostic	t	p
EQS_R				1.198	0.234
	N	45	42		
	Mean	0.7595	0.7225		
	Std. Dev.	0.1450	0.1425		
	Std. Err.	0.0216	0.0220		
EQS_C				-1.156	0.251
	N	45	42		
	Mean	0.5585	0.5998		
	Std. Dev	0.1832	0.1460		
	Std. Err.	0.0273	0.0225		
EQS_A				-0.548	0.585
	N	45	42		
	Mean	0.4967	0.5180		
	Std. Dev	0.1768	0.1846		
	Std. Err.	0.0264	0.0285		

Feedback Specificity and Motivation

Q2a. Do diagnostic feedback recipients become more mastery oriented than their informational feedback counterparts?

As mentioned in the literature review, some research has found a positive correlation between feedback and motivation while others have found no correlation.

To compare changes across the four dimensions of achievement goal orientation, repeated measures were conducted across pre-post (two) administrations of the surveys. Since both the literature and our previous results (Table 4.5) show correlations among the motivation variables, individual repeated measures analyses were conducted, one for each variable being investigated. Results show that the measure mastery goal orientation (MGO) showed a significant change between the pre- and post-surveys, but showed no significant difference between the groups. The prefixes *m* and *f* stand for initial and final surveys, respectively. The MGO * Group interaction is also not significant, which suggests that mastery orientation did not change differently over time between the two groups.

Table 4.12: Differences Between Groups on Mastery Goal Orientation

		Informational	Diagnostic
Initial (m) MGO			
	Mean	3.510	3.447
	Std. Err.	0.107	0.109
Final (f) MGO			
	Mean	3.354	3.160
	Std. Err.	0.116	0.118

Table 4.13: Tests of Within-Subjects Effects for Mastery Goal Orientation

Source	SS	df	MS	F	Sig.	Eta Squared
MGO	**2.335**	**1**	**2.335**	**9.082**	**.003**	**.089**
MGO * GROUP	.204	1	.204	.792	.376	.008
Error(MGO)	23.913	93	.257			

Table 4.14: Tests of Between-Subjects Effects for Mastery Goal Orientation

Source	SS	df	MS	F	Sig.	Eta Squared
Intercept	2154.681	1	2154.681	2277.587	.000	.961
GROUP	.792	1	.792	.837	.363	.009
Error	87.981	93	.946			

Mastery orientation decreased significantly for both feedback types (within-subjects main effect F=9.082, p<0.01, η^2=0.089); however, their MGO scores were not significantly different from each other (between-subjects F=0.837, *n.s.*). This decrease is addressed further in Chapter 5.

Q2b: Do DF recipients become more performance-approach oriented than their IF counterparts?

The results in Tables 4.15-4.17 show that there was no significant main effect for the performance-approach orientation (SHOW; F=3.404, n.s.) and indicates that SHOW did not change significantly during the semester for the full sample. No other changes were significant and no significant differences exist between feedback types.

Table 4.15: Differences Between Groups on Performance Approach (SHOW) Goal Orientation

		Informational	Diagnostic
Initial SHOW			
	Mean	3.533	3.404
	Std. Err.	0.132	0.133
Final SHOW			
	Mean	3.317	3.285
	Std. Err.	0.131	0.132

Table 4.16: Tests of Within-Subjects Effects for Performance-Approach Goal Orientation

Source	SS	df	MS	F	Sig.	Eta Squared
SHOW	1.339	1	1.339	3.404	.068	.035
SHOW * GROUP	.113	1	.113	.287	.593	.003
Error(SHOW)	36.580	93	.393			

Table 4.17: Tests of Between-Subjects Effects for Performance-Approach Goal Orientation

Source	SS	df	MS	F	Sig.	Eta Squared
Intercept	2176.616	1	2176.616	1721.123	.000	.949
GROUP	.306	1	.306	.242	.624	.003
Error	117.612	93	1.265			

Q2c: Do DF recipients become more performance-avoid oriented than their IF counterparts?

Analysis of this research question includes two subscales, AV1 and AV2, each representing a slightly different view of the “avoid embarrassment” goals-driven behavior of students.

There was no significant main effect for AV1 (Tables 4.18-4.20), indicating that there was no significant change in AV1 during the semester for the sample. The lack of significance for AV1 * Group indicates that there is no interaction, which is to say that the slope of the line indicated by each group's avoid orientation over time did not differ. The lack of significance for the main effect of GROUP in the between-subjects results (Table 4.20) indicates that there was no significant difference between groups for the first performance-avoid goal orientation measure.

Table 4.18: Comparison of Means for the first Performance-Avoid (AV1) Measure

		Informational	Diagnostic
Initial AV1		48	47
	Mean	2.250	2.165
	Std. Err.	0.114	0.116
Final AV1			
	Mean	2.204	2.154
	Std. Err.	0.114	0.115

Table 4.19: Tests of Within-Subjects Effects for the first Performance-Avoid (AV1) subscale

Source	SS	df	MS	F	Sig.	Eta Squared
AV1	2.634E-02	1	2.634E-02	.110	.740	.001
AV1 * GROUP	7.916E-03	1	7.916E-03	.033	.856	.000
Error(AV1)	22.184	93	.239			

Table 4.20: Tests of Between-Subjects Effects for the first Performance-Avoid (AV1) subscale

Source	SS	df	MS	F	Sig.	Eta Squared
Intercept	915.884	1	915.884	904.522	.000	.907
GROUP	.248	1	.248	.244	.622	.003
Error	94.168	93	1.013			

One motivation variable showed a significant difference between the groups, but no significant change over the semester: “AV2”, the second of the two performance-avoid subscales. The significant F for Group (F=4.831, p=0.03) indicates that there existed a significant difference in the observed changes of AV2 between the two feedback types during the semester. The recipients of informational feedback began with a higher AV2 mean (measure of concern about what others think about one’s intelligence) and increased in this measure more than the diagnostic feedback recipients (Tables 4.21-4.23).

Table 4.21: Comparison of Means for the second Performance-Avoid (AV2) Measure

		Informational	Diagnostic
Initial AV2		48	47
	Mean	1.948	1.660
	Std. Err.	0.118	0.119
Final AV2			
	Mean	1.969	1.670
	Std. Err.	0.107	0.108

Table 4.22: Tests of Within-Subjects Effects for the second Performance-Avoid (AV2) subscale

Source	SS	df	MS	F	Sig.	Eta Squared
AV2	1.176E-02	1	1.176E-02	.029	.866	.000
AV2 * GROUP	1.234E-03	1	1.234E-03	.003	.956	.000
Error(AV2)	38.362	93	.412			

Table 4.23: Tests of Between-Subjects Effects for the second Performance-Avoid (AV2) subscale

Source	SS	df	MS	F	Sig.	Eta Squared
Intercept	623.500	1	623.500	781.977	.000	.894
GROUP	**4.090**	**1**	**4.090**	**5.129**	**.026**	**.052**
Error	74.152	93	.797			

Feedback Specificity and Epistemological Beliefs

Q3. Do diagnostic feedback recipients become more mature in their epistemological beliefs of chemistry than their informational feedback counterparts?

To compare changes across the three dimensions of epistemological beliefs, repeated measures were conducted across pre-post (two) administrations of the surveys. Recall that questions from Hofer's (1997) Epistemological Beliefs Inventory were employed to measure students' epistemological beliefs of chemistry.

Since both the literature and earlier matrices (Table 4.6) show correlations among the epistemology variables, individual repeated measures analyses were conducted in like manner as the motivational variables analyses. Results (Tables 4.24-4.32) show that only one measure, attainability of truth (TRU), showed a significant change between the pre- and post-surveys, but it showed no significant difference between the groups. There was a significant change in students' epistemological beliefs on the attainability of truth in chemistry, with both groups increasing in their beliefs that truth in chemistry can be obtained ($F=8.032$, $p<0.01$). However, there was not a significant difference between the two groups on this parameter (F=0.949, *n.s.*). That is, the changes that did occur did so independently of feedback type (specificity).

No other statistically significant results were found for the epistemological beliefs measures. Belief in the simplicity and certainty of chemistry knowledge remained stable over time and did not differ between groups. Likewise, there was no significant change in AUTH over time for the sample and no AUTH * GROUP interaction; that is, belief that experts have all the chemistry knowledge did not change differently between feedback types.

Table 4.24: Comparison of Means for Belief in Certainty and Simplicity of Chemistry Knowledge (CERT) subscale

		Informational	Diagnostic
Initial CERT		48	47
	Mean	2.635	2.396
	Std. Err.	0.079	0.080
Final CERT			
	Mean	2.617	2.551
	Std. Err.	0.087	0.088

Table 4.25: Tests of Within-Subjects Effects for the Simple/Certain (CERT) subscale

Source	CERT	SS	df	MS	F	Sig.	Eta Squared
CERT	Linear	.220	1	.220	1.529	.219	.016
CERT * GROUP	Linear	.353	1	.353	2.459	.120	.026
Error(CERT)	Linear	13.363	93	.144			

Table 4.26: Tests of Between-Subjects Effects for the Simple/Certain (CERT) subscale

Source	SS	df	MS	F	Sig.	Eta Squared
Intercept	1235.196	1	1235.196	2372.162	.000	.962
GROUP	1.110	1	1.110	2.132	.148	.022
Error	48.426	93	.521			

Table 4.27: Comparison of Means for Belief in Authority (AUTH) subscale

		Informational	Diagnostic
Initial AUTH		48	47
	Mean	3.156	2.878
	Std. Err.	0.095	0.096
Final AUTH			
	Mean	3.104	3.064
	Std. Err.	0.092	0.093

Table 4.28: Tests of Within-Subjects Effects for the Authority (AUTH) subscale

Source	AUTH	SS	df	MS	F	Sig.	Eta Squared
AUTH	Linear	.213	1	.213	.886	.349	.009
AUTH * GROUP	Linear	.674	1	.674	2.798	.098	.029
Error(AUTH)	Linear	22.402	93	.241			

Table 4.29: Tests of Between-Subjects Effects for Authority (AUTH) subscale

Source	SS	df	MS	F	Sig.	Eta Squared
Intercept	1767.831	1	1767.831	2971.779	.000	.970
GROUP	1.208	1	1.208	2.030	.158	.021
Error	55.323	93	.595			

Table 4.30: Comparison of Means for Belief in the Attainability of Truth (TRU) subscale

		Informational	Diagnostic
Initial TRU		48	47
	Mean	3.170	3.074
	Std. Err.	0.135	0.135
Final TRU			
	Mean	3.340	3.415
	Std. Err.	0.130	0.130

Table 4.31: Tests of Within-Subjects Effects for the Attainability of Truth (TRU) subscale

Source	TRU	SS	df	MS	F	Sig.	Eta Squared
TRU	Linear	3.064	1	3.064	8.032	.006	.080
TRU * GROUP	Linear	.340	1	.340	.892	.347	.010
Error(TRU)	Linear	35.096	92	.381			

Table 4.32: Tests of Between-Subjects Effects for the Attainability of Truth (TRU) subscale

Source	SS	df	MS	F	Sig.	Eta Squared
Intercept	1985.750	1	1985.750	1564.859	.000	.944
GROUP	5.319E-03	1	5.319E-03	.004	.949	.000
Error	116.745	92	1.269			

Q4a. Do overall performance, motivation orientation, or epistemological beliefs of chemistry vary with the number of corrective feedback exposures ("treatment count", TC)?

There is some belief that the beneficial effects of feedback could be compounded over number of treatments. That the study spanned four months and 154 e-quiz questions should provide the opportunity to investigate this perception in the feedback literature. Similarly, the epistemological beliefs literature and motivation orientation literature have demonstrated that change in each of these measures requires time and both external and internal influences, including challenging tasks and a supportive environment (for motivation, see for example, Ames & Archer, 1988; for epistemological beliefs, see for example, Schommer, 1993).

To investigate this question, Pearson product-moment correlations were conducted. The variables included were: treatment count (number of corrective feedback exposures due to missed question, TOT_TC), e-quiz score (EQS_TOT), course score, motivation goal orientation (MGO), performance-approach goal orientation (SHOW), performance-avoidance (avoid embarrassment) orientation (AV1 & AV2), certain/simple views (CERT), authority dimension (AUTH), and attainability of truth (TRU). The prefix "f" denotes that these measures were taken from the end-of-semester survey, or post-survey. The correlation table is shown in Table 4.33.

There is a significant positive correlation between the number of corrective feedback exposures (TOT_TC) and course score for each group. No other correlations were significant. Performance as measured by course score seems to vary with the number of corrective feedback exposures, such that an increase in treatment count corresponds with an increase in overall course score. The results likewise suggest that an increase in number of feedback treatments does not necessarily increase motivation or beliefs measures included in this study.

The breakdown of these relationships by feedback type is the focus of the next pair of questions.

Q4b. What is the relationship among treatment count, motivation, and epistemological beliefs of chemistry for the informational feedback recipients?

Q4c: What is the relationship among treatment count, motivation, and epistemological beliefs of chemistry for the diagnostic feedback recipients?

These questions differ from the previous questions in that they provides a different vantage point. Other researchers have noted relationships between specificity of feedback and motivation (Ammons, 1956; Butler & Nisan, 1986; Earley, et al, 1990; Johnson, et al, 1993), and between motivation and epistemological beliefs (Nicholls, et al, 1984, 1990), but Ammons' statement that more specific feedback yields higher motivation and performance has been challenged on several occasions, the most notable being the meta-study by Kluger & DeNisi (1996) and references therein.

A new correlation analysis was conducted per feedback type that included all of the variables in the study. The full-sample correlation matrix is shown in Table 4.33. Group results for each feedback type are shown in Table 4.34 for the informational feedback group and Table 4.35 for the diagnostic feedback group. The results show a significant linear correlation between treatment count and final course score for the diagnostic feedback group, but surprisingly, not for the informational feedback group. Transversely, there appears to exist a significant linear correlation between final course score and final mastery goal orientation measures for the informational feedback group, but not the diagnostic feedback group.

Table 4.33: Correlations for All Students in Sample

		TOT_TC	EQS_TOT	course score	FMGO	FSHOW	FAV1	FAV2	FCERT	FAUTH	FTRU
TOT_TC	Corr.	1.000									
	Sig.	.									
	N	95									
EQS_TOT	Corr.	-.054	1.000								
	Sig.	.622	.								
	N	87	87								
course score	Corr.	.221	.440	1.000							
	Sig.	.032	.000	.							
	N	95	87	95							
FMGO	Corr.	-.021	.336	.297	1.000						
	Sig.	.843	.001	.003	.						
	N	95	87	95	95						
FSHOW	Corr.	.028	.110	.152	.000	1.000					
	Sig.	.784	.311	.143	.997	.					
	N	95	87	95	95	95					
FAV1	Corr.	-.110	.023	-.057	.198	.458	1.000				
	Sig.	.287	.835	.586	.055	.000	.				
	N	95	87	95	95	95	95				
FAV2	Corr.	-.155	.063	-.007	.154	.305	.616	1.000			
	Sig.	.134	.561	.944	.137	.003	.000	.			
	N	95	87	95	95	95	95	95			
FCERT	Corr.	.199	-.170	-.132	-.125	-.037	.116	.040	1.000		
	Sig.	.053	.115	.203	.228	.725	.265	.702	.		
	N	95	87	95	95	95	95	95	95		
FAUTH	Corr.	.161	.027	.010	-.151	.078	.065	-.004	.493	1.000	
	Sig.	.119	.803	.924	.143	.450	.532	.967	.000	.	
	N	95	87	95	95	95	95	95	95	95	
FTRU	Corr.	-.055	-.043	-.037	.140	.042	.022	-.044	.010	.095	1.000
	Sig.	.601	.693	.725	.179	.690	.837	.673	.922	.365	.
	N	94	86	94	94	94	94	94	94	94	94

Table 4.34: Correlation Matrix for Students Receiving Informational Feedback

		TOT_TC	EQS_TOT	course score	FMGO	FSHOW	FAV1	FAV2	FCERT	FAUTH	FTRU
TOT_TC	Corr.	1.000									
	Sig.	.									
	N	48									
EQS_TOT	Corr.	-.152	1.000								
	Sig.	.320	.								
	N	45	45								
course score	Corr.	.149	.429	1.000							
	Sig.	.312	.003	.							
	N	48	45	48							
FMGO	Corr.	-.197	.395	**.452**	1.000						
	Sig.	.179	.007	**.001**	.						
	N	48	45	**48**	48						
FSHOW	Corr.	.040	.156	.144	-.212	1.000					
	Sig.	.786	.306	.330	.149	.					
	N	48	45	48	48	48					
FAV1	Corr.	-.202	.063	-.028	-.066	.463	1.000				
	Sig.	.168	.681	.848	.656	.001	.				
	N	48	45	48	48	48	48				
FAV2	Corr.	-.125	.155	-.014	.033	.293	.607	1.000			
	Sig.	.397	.309	.924	.823	.043	.000	.			
	N	48	45	48	48	48	48	48			
FCERT	Corr.	.386	-.209	-.128	-.192	.042	.126	.047	1.000		
	Sig.	.007	.167	.385	.190	.775	.392	.750	.		
	N	48	45	48	48	48	48	48	48		
FAUTH	Corr.	.058	-.010	-.039	-.065	.173	.252	.022	.520	1.000	
	Sig.	.693	.950	.791	.661	.239	.084	.883	.000	.	
	N	48	45	48	48	48	48	48	48	48	
FTRU	Corr.	-.147	.024	-.034	.043	-.110	-.040	-.151	-.098	.187	1.000
	Sig.	.325	.875	.821	.776	.461	.788	.312	.510	.209	.
	N	47	44	47	47	47	47	47	47	47	47

Table 4.35: Correlations Matrix for Diagnostic Feedback Recipients

		TOT_TC	EQS_TOT	course score	FMGO	FSHOW	FAV1	FAV2	FCERT	FAUTH	FTRU
TOT_TC	Corr.	1.000									
	Sig.	.									
	N	47									
EQS_TOT	Corr.	.051	1.000								
	Sig.	.749	.								
	N	42	42								
course score	Corr.	**.344**	.455	1.000							
	Sig.	**.018**	.002	.							
	N	**47**	42	47							
FMGO	Corr.	.109	.277	.157	1.000						
	Sig.	.464	.075	.293	.						
	N	47	42	47	47						
FSHOW	Corr.	.010	.063	.165	.181	1.000					
	Sig.	.945	.693	.267	.223	.					
	N	47	42	47	47	47					
FAV1	Corr.	-.005	-.030	-.095	.484	.457	1.000				
	Sig.	.975	.850	.527	.001	.001	.				
	N	47	42	47	47	47	47				
FAV2	Corr.	-.279	-.053	.025	.233	.326	.642	1.000			
	Sig.	.057	.737	.867	.115	.025	.000	.			
	N	47	42	47	47	47	47	47			
FCERT	Corr.	-.029	-.134	-.132	-.079	-.121	.098	.008	1.000		
	Sig.	.848	.397	.378	.598	.419	.512	.956	.		
	N	47	42	47	47	47	47	47	47		
FAUTH	Corr.	.281	.062	.081	-.249	-.028	-.197	-.052	.460	1.000	
	Sig.	.056	.697	.590	.091	.852	.185	.727	.001	.	
	N	47	42	47	47	47	47	47	47	47	
FTRU	Corr.	.056	-.108	-.046	.234	.187	.098	.083	.125	.000	1.000
	Sig.	.708	.494	.761	.113	.207	.514	.578	.402	.999	.
	N	47	42	47	47	47	47	47	47	47	47

These results were unexpected and surprising, so scatter plots were created. Scatter plots for mastery goal orientation (MGO) and treatment count (TC) each with course score revealed a curvilinear relationship with course score for each feedback type (see Figures 4.1.1-2 and Figures 4.2.1-2).

Curvilinear regression analysis revealed significant quadratic relationships between treatment count and course score for both informational feedback type ($F_{(2,45)}=14.33$, $p<0.001$) and diagnostic feedback type ($F_{(2,44)}=3.72$, $p<0.05$). Because visual inspection of the scatter plot for the diagnostic feedback sample showed an inversion at very low TC values, a cubic regression analysis was conducted. For purposes of comparison, this analysis was executed for both feedback samples. The result was an even better fit of the model with a cubic regression of treatment count on course score for the diagnostic feedback type ($F_{(3,43)}=5.08$, $p<0.01$) and the cubic model fit was still statistically significant for the informational feedback type ($F_{(3,44)}=9.74$, $p<0.001$). These results are shown in Tables 4.36 & 4.37.

Figure 4.1.1: Scatter Plot of TC with Course Score for Informational Feedback Type

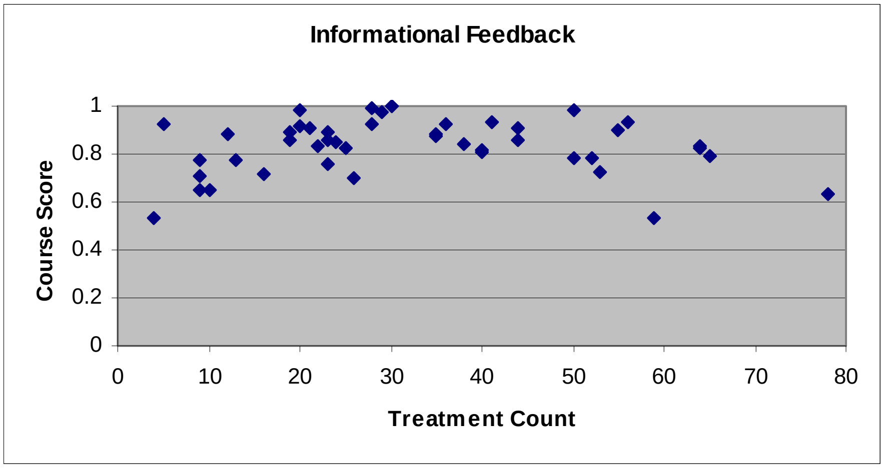


Figure 4.1.2: Scatter Plot of TC with Course Score for Diagnostic Feedback Type

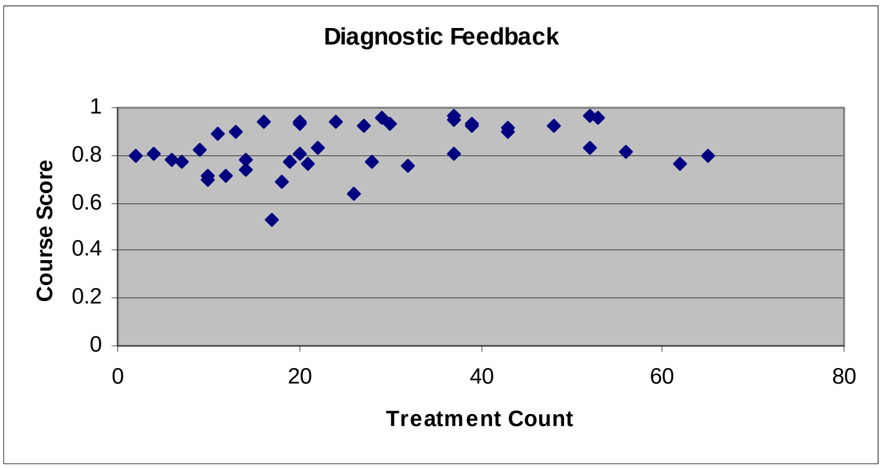


Figure 4.2.1: Scatter Plot of Course Score with Mastery Orientation Score for Informational Feedback Type

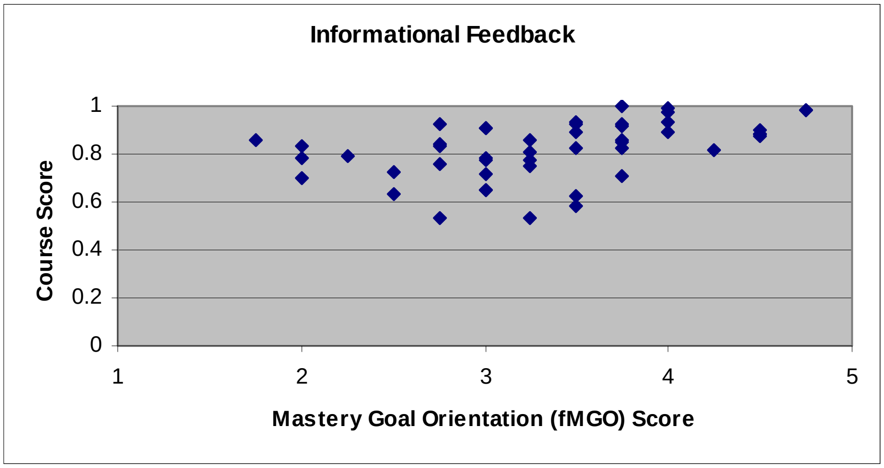


Figure 4.2.2: Scatter Plot of Course Score with Mastery Orientation Score for Diagnostic Feedback Type

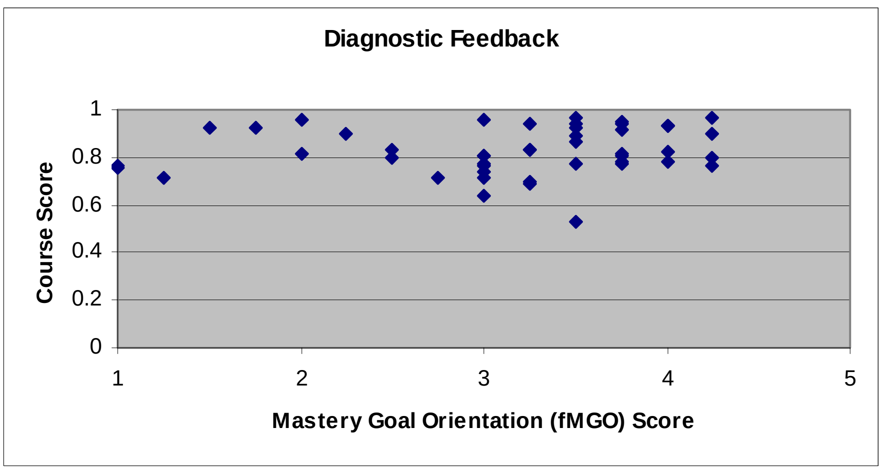


Table 4.36: Values of R, R^2, and adjusted R^2 for quadratic regressions

Independent Variable	Dependent Variable	R	R^2	adjusted R^2	Std. Err.
					Informational (IF)
TC	Course score	0.62	0.39	0.36	0.095
fMGO	Course score	0.51	0.26	0.23	0.11
					Diagnostic (DF)
TC	Course score	0.38	0.14	0.11	0.094
fMGO	Course score	n.s.	n.s.	n.s.	n.s.

Table 4.37: Values of R, R^2, and adjusted R^2 for cubic regressions

Independent Variable	Dependent Variable	R	R^2	adjusted R^2	Std. Err.
					Informational (IF)
TC	Course score	0.63	0.40	0.36	0.096
fMGO	Course score	0.54	0.29	0.24	0.10
					Diagnostic (DF)
TC	Course score	0.51	0.26	0.21	0.088
fMGO	Course score	n.s.	n.s.	n.s.	n.s.

Similarly, the regression of mastery goal orientation score from the post-survey (fMGO) showed improved fit in the quadratic model for the informational feedback type ($F_{(2,45)}=7.84$, $p<0.05$), signaling that the linear trend was insufficient. The curvilinear models were not a good fit for the diagnostic feedback scores. Since fMGO seems to correlate with number of treatments (TC) on one feedback type selectively, this suggests a possible moderating effect of fMGO with feedback type and the number of treatments. However, according to Baron & Kenny (1986), there must be an fMGO * Type interaction in order to perform tests for moderating effects. By changing fMGO to a dichotomous variable, a 2x2 ANOVA was conducted according to Baron & Kenny's procedures and the interaction was not found (F=0.032, *n.s.*). Using the pre-test MGO score as a covariate did not change these results.

It is possible to test for moderating or mediating effects by linear regression if the linear transformation may be computed (Cohen & Cohen, 1983). For example, to conduct regression analyses on a logarithmic function, one may transform it into a linear function by computing the inverse log of the original function. However, in the case of this study there is not enough known presently in the literature about the relationship between mastery goal orientation and performance to formulate a reasonable transformation for the data. That is, several functions can produce a nonlinear relationship between two variables, but choosing the appropriate one requires knowledge of not only the order of the relationship, but also the influence of a constant, c—for example, deciding between a power function in $y=x^c$ versus $y=cx^3$ requires knowledge of how c relates with y. Since the literature presently appears to assume a linear relationship, such a transformation may be unnecessary. Until the curvilinear relationship found here is replicated under more controlled conditions, we cannot determine whether the observed relationship is the result of measured subscales or other measures not taken. Even so, we can still conduct some analyses based on trends observed in the plots of the data.

One problem with the plots involving treatment count (TC) is that the TC measure is confounded: it has participation and number of e-quizzes hidden in it. For example, the student who missed 6 e-quiz questions (TC=6) could have done so in one e-quiz or across all 15 e-quizzes, spanning the nine weeks during which they were offered. Since the e-quizzes were voluntary in nature and since “treatment” occurred only for incorrect responses, treatment did not occur at a standard rate. Other plots were created to help draw out the story behind the data. To begin, plots of course score with e-quiz scores were created for each feedback type. The e-quiz score is the ratio of the number of questions correctly answered to the number of questions attempted. The e-quiz score offers a view of “grades” that allows us to determine whether students who made A’s on the e-quizzes also made A’s in the class, reflecting an interpretation of the correlation results given previously (Table 4.33). Recall that the course score used for analyses did not include any contribution from the e-quiz performance. The trend seen in the plots (Figures 4.3.1-2) confirm that students who made A’s on the e-quizzes also made A’s in the course, but shows that these are not the only students who made an A in the course. There is a great range of e-quiz scores that corresponded to A’s in the course, regardless of feedback type.

Figure 4.3.1: Course Grade by E-quiz Scores for Informational Feedback Recipients

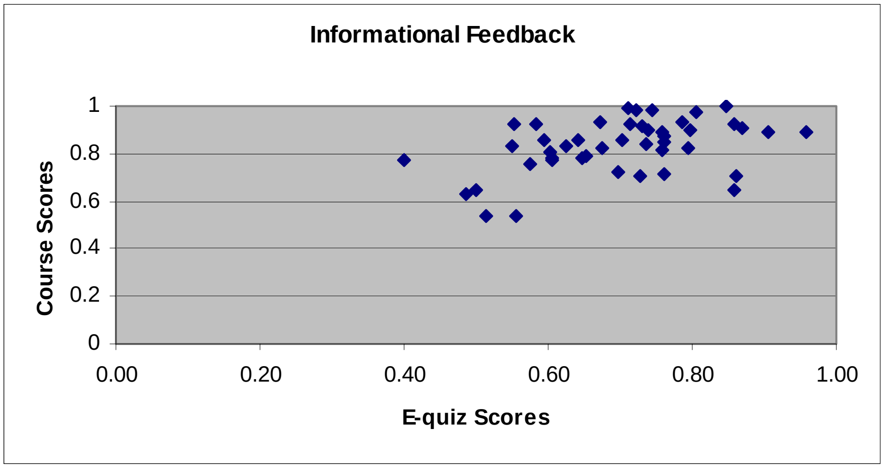
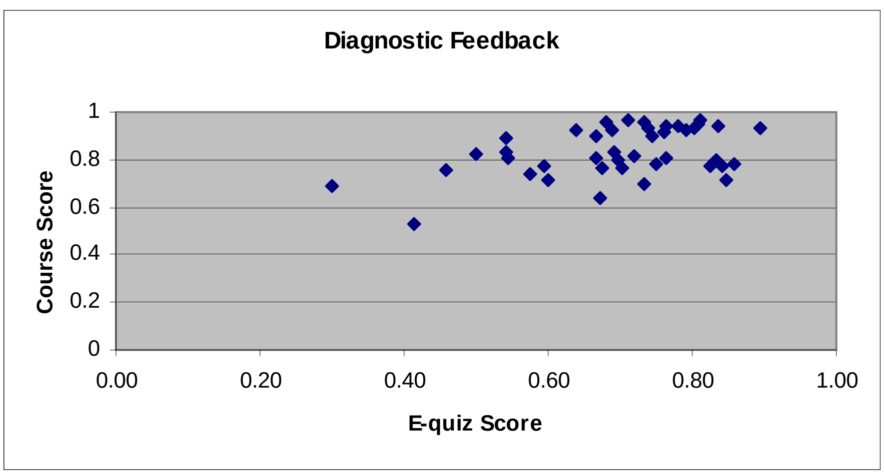


Figure 4.3.2: Course Grade by E-quiz Scores for Diagnostic Feedback Recipients



No differential trend between feedback types was found in these plots. However, the e-quiz scores do not reveal differences by participation. For example, a student receiving an e-quiz score of 0.80 (reflecting a B letter grade) may have done so by taking a single e-quiz or may have taken all e-quizzes. These plots do not distinguish between those two hypothetical students. Thus, two additional sets of plots were created. One set uses an average treatment count (avgTC) as an attempt to better normalize treatment count by computing the number of treatments (questions missed) divided by the number of e-quizzes taken. These plots (Figures 4.4.1-2) show a general downward trend with increasing average TC, which is interpreted that, in general, students in either feedback group who received more treatments per e-quiz showed a decreased corresponding course score. This makes sense when treatments are interpreted as the number of questions missed; students who missed more questions on the assignment also received a lower course score (even though the assignment score was not included in the course score computation).

No differential trend between feedback types was found in these plots. Students of different levels of accomplishment show no real differences between feedback types. But these plots do not provide information as to participation. That is, since participation is effectively voluntary, it would be interesting to know whether, for example, A-letter grade level students and F-level students opted out of the e-quizzes differentially between the two feedback types. That is, did A-level students start taking the e-quizzes, then get tired of the diagnostic feedback and stop taking them where the informational feedback A-level students did not? Did this change occur for any other subset of students, for example, the D or F-level students who tend to struggle through the course? Such a trend could help account for the curvilinear trends described previously.

To examine this possibility, a set of plots showing e-quiz participation was created. One set shows the participation of students in the "Option 1" e-quizzes while the other set shows the participation of students in the "Option 2" e-quizzes. Each week, two e-quizzes were offered, Option 1 and Option 2. In order to qualify for the "extra" points at the end of the semester, students had to "pass" an e-quiz. A "pass" meant that the student answered at least two-thirds of the questions correct for at least one of the two e-quizzes offered each week. Passing two e-quizzes was redundant with regard to the "extra" points.

The student participation plots, which resemble gel electrophoresis results, show the spread of student course scores over the various e-quizzes (Figures 4.5.1-2 and Figures 4.6.1-2). Each "line" represents one student who took that e-quiz and received the corresponding course score. The columns of lines represent individual e-quizzes. Note that an Option 2 e-quiz did not exist for the seventh e-quiz week.

The plots show no difference in trends over time or between feedback types. The A and B grade level students did not decrease participation over time selectively in either feedback treatment. Likewise, the C, D, and F-level students did not show a trend beyond the reduced participation compared with that of the A and B course grade students. The result these plots show is not surprising: students who participate more in the class work also tend to make higher scores in the course.

The reason for the obvious break in the plot for the diagnostic feedback group is unknown. Within the diagnostic feedback group, there is an absence of students who received a mid-B grade in the course. A similar break is observed in the informational feedback plot for mid-A grades. The reason behind these breaks is unknown.

In order to further explore the participation aspect of the e-quizzes, a plot of course score by number of e-quizzes taken was created for each feedback type. The expected result was that greater participation would be related to higher course scores. However, this was not seen. The plots revealed significant cubic relationships for each feedback type, but the nature of the curves differed. (Figures 4.7.1-2) The most notable difference is that course score dips immediately for students assigned to the diagnostic feedback treatment, but increases for those students receiving informational feedback. At this point, it is unknown whether the dip noticed for the diagnostic feedback type (Figure 4.7.2) is due to lower-scoring students receiving the diagnostic feedback type selectively taking only a few e-quizzes while mid-scoring students receiving the informational feedback type selectively took the same number of e-quizzes or whether e-quiz score had anything to do with the relationship seen.

In order to get a better look at these differences, a series of plots could be constructed of course score versus average treatment count for students participating in 2 or more, 3 or more, ….and on up to 12 or more e-quizzes, with the intention to determine at what point, if any, a change in the pattern could be observed. However, in order to be reliable, such a process would require greater numbers of participants than we have in our study. This line of questioning is thus left to future research. Based on others' suggestions that new feedback requires some time for students to become acquainted with it and learn to utilize it (Annett, 1969), we are not altogether surprised that students receiving diagnostic feedback would perform less well initially than their informational feedback peers, until the time in which the diagnostic feedback recipients became familiar with the new feedback style. Confirmation of this hypothesis is left to future research.

Figure 4.4.1: Scatter Plot of Course Score with Average TC for Informational Feedback Type

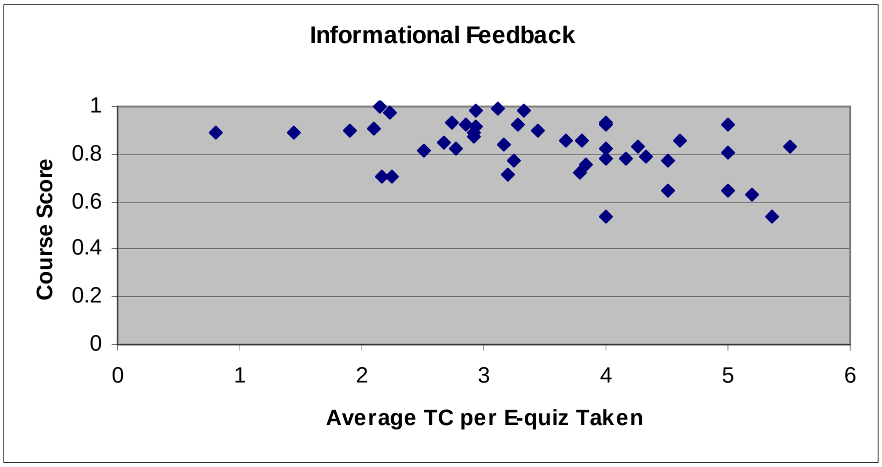
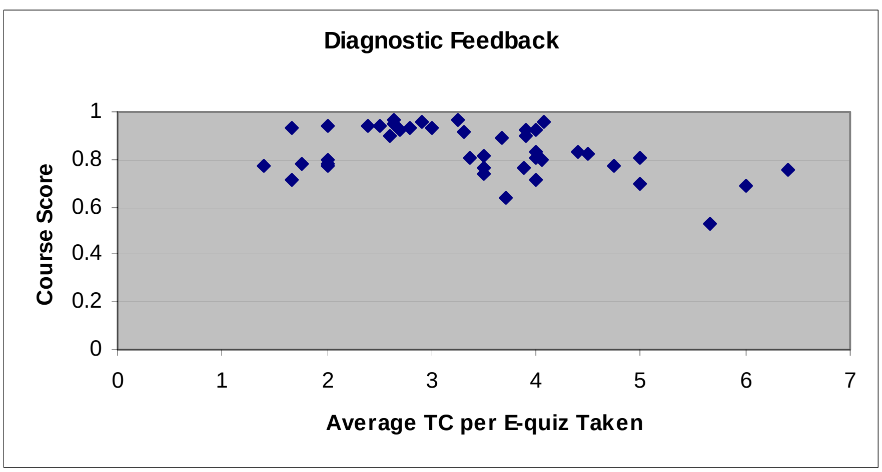
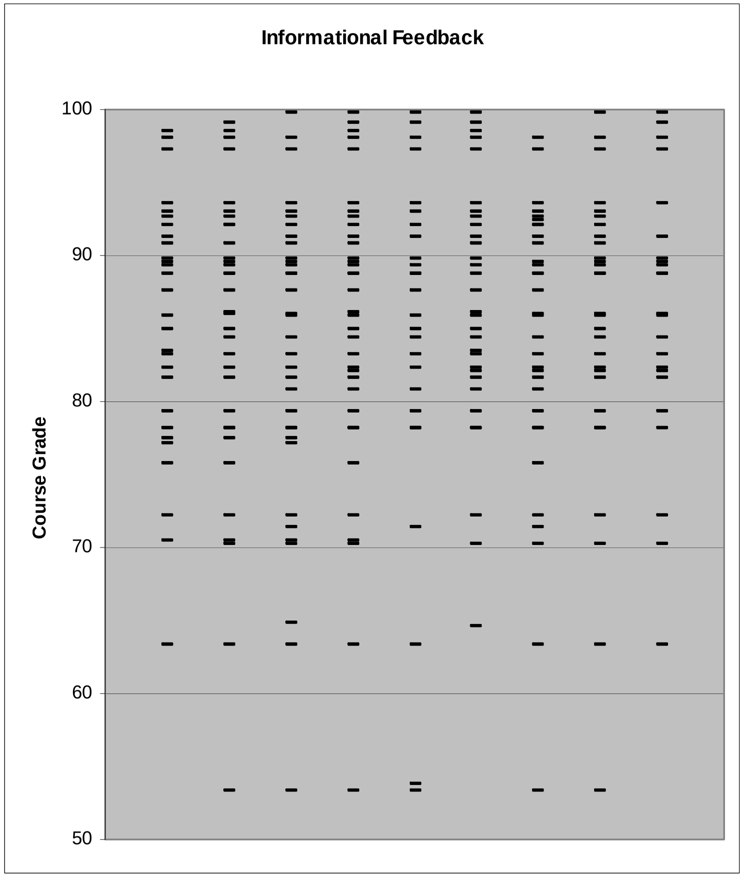
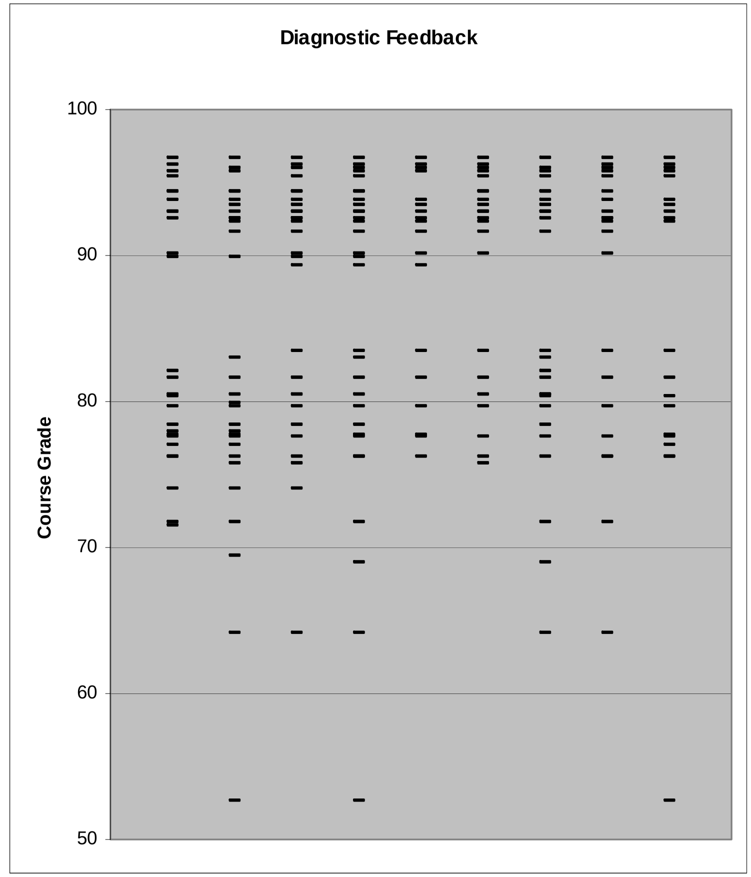


Figure 4.4.2: Scatter Plot of Course Score with Average TC for Diagnostic Feedback Type

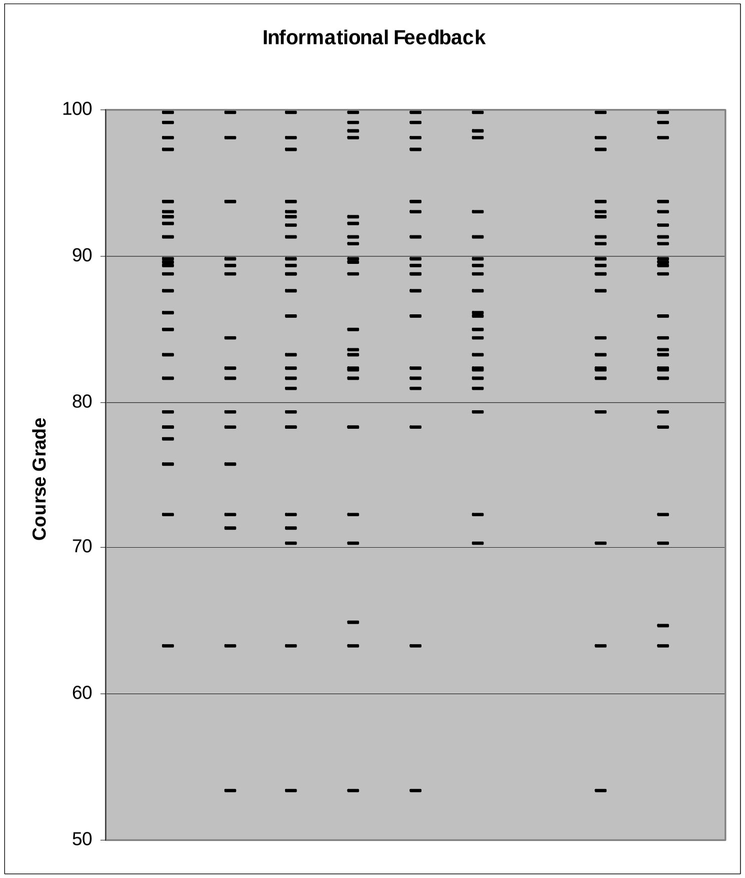


Figures 4.5.1 and 4.5.2: Course Score with Participation in “Option 1” E-quizzes for Each Feedback Type





Figures 4.6.1 and 4.6.2: Course Score with Participation in “Option 2” E-quizzes for Each Feedback Type



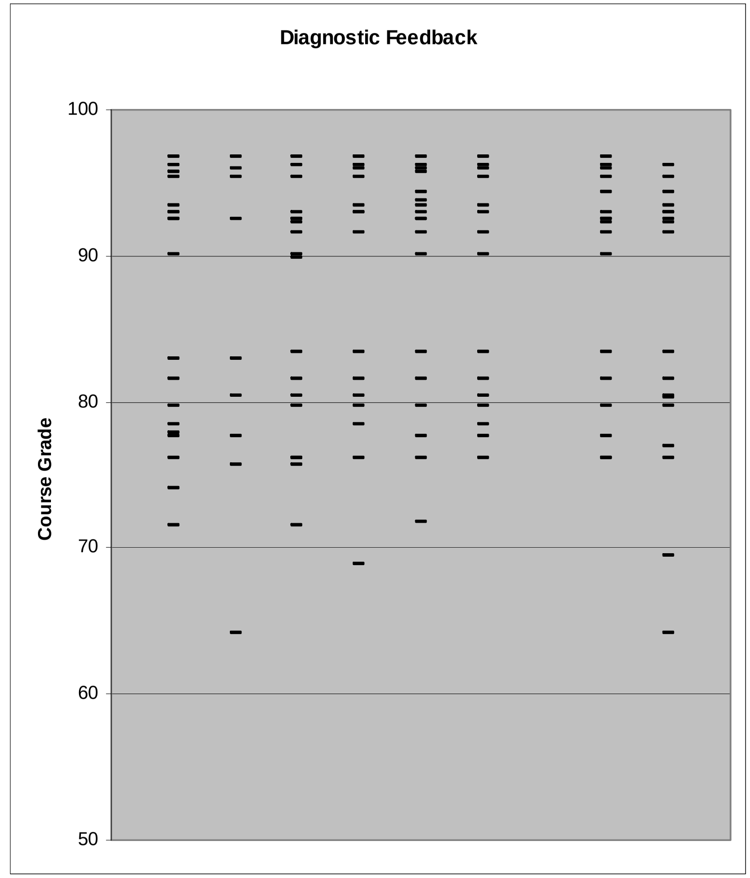


Figure 4.7.1: Course Score with Total Number of E-quizzes Taken for Informational Feedback Type

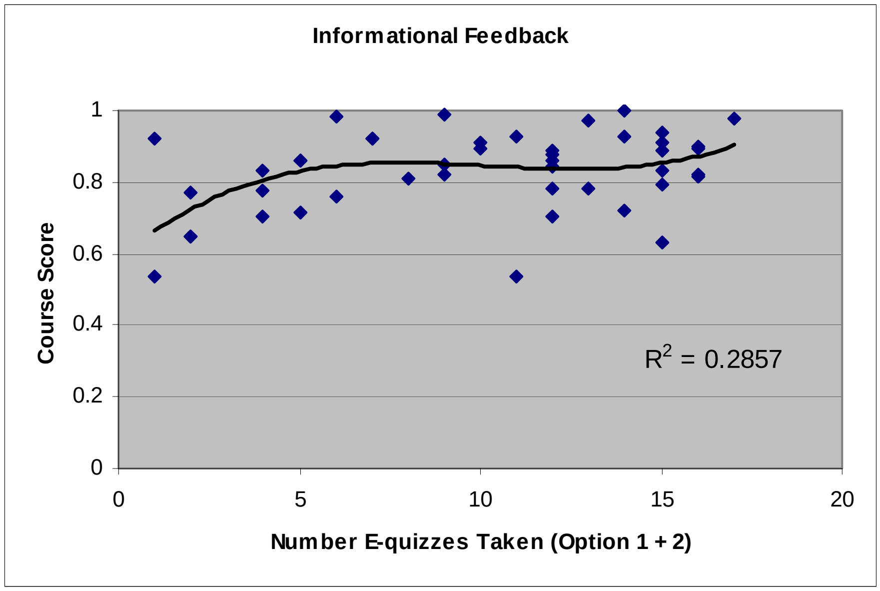
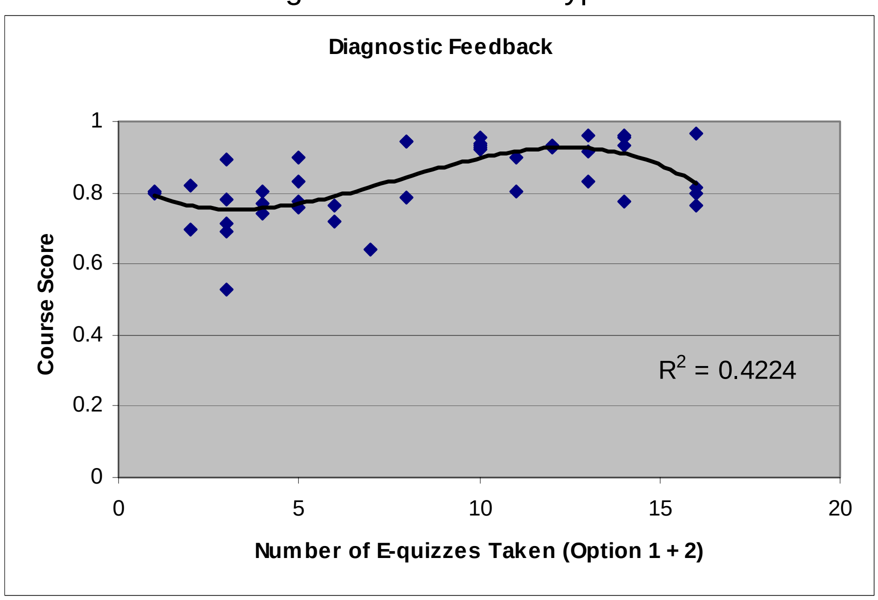


Figure 4.7.2: Course Score with Total Number of E-quizzes Taken for Diagnostic Feedback Type



Quasi-qualitative Analysis

Q5: Do students who receive diagnostic feedback outperform their informational feedback classmates on a same or similar question posed (1) soon after initial exposure & correction and (2) later, on an examination?

We wish to discover whether there is a difference between feedback types for individual students that perhaps was not perceived in the averages and statistical measures employed previously. In particular, this analysis investigates three case studies in detail to compare success rates of students on questions which they have previously encountered and on which they have received feedback. To help with this analysis, a few e-quiz questions were purposely repeated on either the same e-quiz or a close-in-time e-quiz to determine if a difference exists between the two feedback types, as such a finding would suggest differential effects on retention and transfer of information. On one serendipitous occasion the course professor included on an examination one of these questions and two closely related problems. Thus, this analysis looks at the relative success rates of the students in the two feedback groups on these three questions.

In-depth analysis similar to qualitative analysis was employed for this question, comparing each student's performance across three viewings of the same or a related problem. The results show no significant difference between the feedback types, but other interesting findings emerged.

Case Study 1

The text for the problem in the first case study is shown in Figure 4.8.

Figure 4.8: Case Study 1

Case Study 1: Conceptual Buffer Problem (All 3 viewings)
A buffer is composed of NH_3 and NH_4Cl. How would this buffer solution control the pH of a solution when a small amount of a strong base is added?

A. The OH^- reacts with the NH_4^+.
B. The OH^- reacts with the Cl^-.
C. The OH^- reacts with the NH_3.
D. The OH^- reacts with the H_2O.

This problem was viewed verbatim three times by the student sample: twice on the same e-quiz and once on the third examination. The feedback text is provided in the Appendix A.3. The correct answer to this problem is option A. A great majority of the students answered this question correctly on the first try. For the 5 informational feedback recipients and 1 diagnostic feedback recipients who missed the problem on the first viewing, 3 and 1 of them, respectively, correctly answered it on the second attempt a few questions later on the same e-quiz. That leaves 2 students in the informational feedback group who missed the problem on the first attempt, then missed it again. It is notable that the students had to read the feedback in order to deduce the correct response: the answer was not clearly listed as the first line of feedback, so we wonder whether the students even read the feedback. In addition, the multiple choices were randomized, so the order of the responses for the second viewing was very likely different.

In regard to examination performance, 2 of the initial 5 informational feedback students missed the exact same question on the examination. The single diagnostic-feedback recipient who answered the question wrong at the first viewing and correct at the second viewing later answered the examination question correctly.

Obviously, information gained from this type of descriptive analysis is limited when students perform very well, so next are more challenging cases.

Case Study 2

The second case study looks at a more challenging problem that appeared on the e-quiz the following week after Case Study 1. This problem involves finding the reaction rate for a multi-step reaction. The complexity of the problem is additionally increased by the use of abstract letters in place of actual chemical elements. The correct answer to both questions shown in Figure 4.9 is A.

Figure 4.9: Case Study 2

Case Study 2: Reaction Rate (1st and 2nd Viewings, on e-quizzes)

Consider the multistep reaction that has the overall reaction
2A + 2B --> C + D.
What is the rate law expression that would correspond to the following proposed mechanism?

A + B --> I (fast)
I + B --> C + X (slow)
X + A --> D (fast)

A. Rate = $k[A][B]^2$
B. Rate = k[I][B]
C. Rate = k[A][B][I]
D. Rate = k[A][B][I][X][D]

Case Study 2: Reaction Rate (3rd Viewing, on exam)

The overall reaction
$NO_2(g) + CO_2(g) \rightarrow NO(g) + CO_2(g)$
Is proposed to proceed by the mechanism
$2\ NO_2(g) \rightarrow NO_3(g) + NO(g)$ slow
$NO_3(g) + CO(g) \rightarrow NO_2(g) + CO_2(g)$ fast
Which of the following rate laws would be consistent with this mechanism?

A. rate = $k[NO_2]^2$
B. rate = $k[NO_2][CO]$
C. rate = $k[NO_3][CO]$
D. rate = $\frac{[NO][CO_2]}{[NO_2][CO]}$
E. None of the rate laws given here

Figure 4.10: Case Study 3

Case Study 3: pH of a Salt (1st viewing, on e-quiz)

A question states: What is the pH of a 0.12 M solution of anilinium nitrate, $C_6H_5NH_3NO_3$? K_b for aniline ($C_6H_5NH_2$) is 4.2 x 10^{-10}.
Will the solution be acidic, basic, or neutral? Or is it predictable at all?

A. acidic
B. basic
C. neutral
D. cannot be predicted from the information given

Continuing the question: What is the pH of a 0.12 M solution of anilinium nitrate, $C_6H_5NH_3NO_3$? K_b for aniline ($C_6H_5NH_2$) is 4.2 x 10^{-10}.
It is apparent that we must determine $[H_3O^+]$ in order to arrive at pH (that's one way to tackle this problem). Which reaction will lead us to that concentration?

A. $C_6H_5NH_3^+ + H_2O \leftarrow\rightarrow C_6H_5NH_2 + H_3O^+$
B. $HNO_3 + H_2O \leftarrow\rightarrow NO_3^- + H_3O^+$
C. $C_6H_5NH_3NO_3 \leftarrow\rightarrow C_6H_5NH_3^+ + NO_3^-$
D. $2\ H_2O \leftarrow\rightarrow H_3O^+ + OH^-$
E. $C_6H_5NH_3NO_3 + H_2O \leftarrow\rightarrow C_6H_5NH_3NO_3 + H_3O^+$

Now let's finish out the question. What is the pH of a 0.12 M solution of anilinium nitrate, $C_6H_5NH_3NO_3$? K_b for aniline ($C_6H_5NH_2$) is 4.2 x 10^{-10}.

A. 2.77
B. 5.15
C. 11.23
D. None of these
E. Unable to determine.

Case Study 3: pH of a Salt (2nd viewing, on e-quiz)

A question asks:

What is the pH of a 0.12 M solution of ammonium nitrate, NH_4NO_3. K_b for ammonia, NH_3, is 1.8×10^{-5} .

Will the pH be acidic, basic, neutral, or cannot be determined?

A. pH < 7
B. pH > 7
C. pH = 7
D. cannot be determined from the information given

A question asks:

What is the pH of a 0.12 M solution of ammonium nitrate, NH_4NO_3. Kb for ammonia, NH_3, is 1.8×10^{-5}.

What is the K expression to use to calculate [H_3O+] directly? (Recall that brackets [ ] denote concentration.)

A. K_a = [NH_3][H_3O+] / [NH4+]
B. Kb = [HNO_3][OH-] / [NO_3-]
C. Kb = [NH_4+][OH-] / [NH_3]
D. Kb = [NH_3][HNO_3] / [NH_4NO_3]
E. None of these

A question asks:

What is the pH of a 0.12 M solution of ammonium nitrate, NH_4NO_3. Kb for ammonia, NH_3, is 1.8×10^{-5}.

Calculate the pH of the solution.

A. 5.09
B. 2.83
C. 7.00
D. 11.2

Case Study 3: pH of a Salt (3rd Viewing, on exam)

What is the pH in a solution made by dissolving 0.100 mole of sodium acetate $NaCH_3COO$ in enough water to make one liter of solution?

K_a for CH_3COOH is 1.80×10^{-5}.

A. 5.56×10^{-10}
B. 5.56×10^{-11}
C. 7.46×10^{-6}
D. 1.80×10^{-6}
E. 1.34×10^{-9}
F. 9.25
G. 10.25
H. 5.13
I. 5.74
J. 8.87 (correct)

Of the 16 informational- and 14 diagnostic-feedback students who responded incorrectly the first try, 9 and 11, respectively, correctly answered it the second time they saw it—three questions later on the same e-quiz. The remaining 7 and 3 students missed it again, even though feedback was given just moments beforehand. It was surprising that more students didn't answer it correctly. This finding generated question 6 concerning whether students learned from the feedback.

Case Study 3

The analysis of this comparison is different from the previous case studies because it is a very challenging problem for students (Figure 4.10), so it was split into three problems for each e-quiz. Also, the first two viewings occurred across two different e-quizzes given the same week and the examination question involved a different type of salt than the practice questions, so that a moderately different procedure was needed to solve the problem. Thus, the analyses look across feedback type, retention, and transfer, though one must keep in mind that this is one single case and so may not be representative of a sample of similar questions.

Beginning with analyses across the two e-quizzes, findings will be reported as before, listing first the number of students in the informational feedback group followed by those in the diagnostic feedback group. For part A of the question, which involved predicting the pH of the salt solution, 7 and 6 students, respectively, missed it the first viewing, then 4 and 2, respectively, got it correct the second try. Next, part B of the two questions are different, but are still “middle steps” of very similar problems, and so serve as a check of whether the student can successfully complete those middle steps. Of the 10 and 8 students, respectively, who missed the first viewing part B, 7 and 4 of them, respectively, correctly answered part B in the second viewing. And for part C, in which the student had to use the prior steps and given information to calculate pH, 12 and 9 students, respectively, missed it and of those, 6 and 2, respectively, correctly answered the question on the other e-quiz.

To help track performance across the many levels of this question and to the examination performance, a coding system was invented to describe the changes from one viewing to another. The code “null” was assigned to students who got the question right on two subsequent viewings: a correct, correct sequence of responses by the student. Code “neg” was assigned for a correct, incorrect sequence. Code “pos” was assigned for an incorrect, correct sequence. Simultaneous comparison across three viewings is possible by the combination of two codes, with one adjustment: since retention of feedback information would be viewed as a positive result, carry over of a correct answer from view 2 to view 3 is coded as positive. Thus, the only “null” codes are reserved for those students who never missed the e-quiz question and therefore did not get correctional treatment. For example, the “best” case scenario in support of feedback would be a pos-pos code combination because that reflects the student who missed the question on the first try then got it right on the second and third viewings. These codes were then compared across the levels of steps in each e-quiz question and across each viewing of the e-quiz and the examination to arrive at an estimate of feedback-induced improvement for each feedback type.

Table 4.38: In-Depth Examination of Relative Success Across Three Different, but Related, Problems

	# of pos-pos	#of neg-pos	# of null-pos	(blank) pos	of possible
Info	4	1	1	2	25
Diag	1	1	1	1	18

The tallies in Table 4.18 show that of the 25 students in the informational feedback group who received treatment at some point during practice, a total of 8 of them answered the related examination question correctly. Of the 18 students who received diagnostic treatment at some point during e-quiz practice, 4 of them answered the related examination question correctly. Also notable is that 4 and 1 students, respectively, needed only one treatment before answering the question correct on the examination: they appear to have learned from a single mistake, carried that knowledge over to a similar problem on the next e-quiz, and then transferred that knowledge (plus any additional studying) to the most difficult problem on the examination.

Q6: Do students who receive feedback reproduce their mistakes? Are diagnostic feedback recipients less likely than their informational feedback counterparts to make the same mistake again?

This post-hoc question continues the analysis from Question 5, but looks at it in a slightly different way. This question analyses the content of the actual response the student chose in the first viewing, then checks whether the student replicated that mistake in subsequent viewings. In examining rates of success for each quiz question above, it appeared as though some students were replicating their mistakes. However, in looking at each individual student's responses to the three cases examined in Question 5, a different story unfolded. No student made the same mistake twice in two of the three cases examined. The students did not know ahead of time when or if they would be given a second viewing, so it stands to reason that the students paid at least enough attention to note their previous responses and note that they were incorrect.

But there's more to report. Based on the same three cases studied in Question 5, the individual response data show interesting behaviors in responses to the questions. Results will be presented in pairs, with the number of students in the informational group followed by those in the diagnostic group. In Case 2 (reaction rate problem), of the 9 and 12 students, respectively, who answered option B (see Table 4.10) in the first viewing, 7 and 8 of them, respectively, subsequently answered the question correctly. The remaining 2 and 4 of them, respectively, answered option 3 in the second viewing. This change is a step-wise change. That is, there are many steps in this problem and after missing the question the first time the student seems to have corrected his thinking on that one step; however, in the next encounter with the question, these students made a mistake in the next step of the solution, which involved spectator ions. Every one of these six students who missed the problem the second time missed it in this way. The other students got the problem correct the second try. This finding brings to mind a quote by Dorothy Gabel, a researcher in chemical education misconceptions:

> "Many chemistry concepts are very abstract. If there is nothing in long-term memory to which a new concept can be

> related, then it will either not be stored, or it will be stored as a single entity. Hence, if something does exist to which the new concept can be related, then learning occurs…"

In studying trends in the responses of Case Study 3, it was interesting that the case in which the students repeated their mistake was here, in what may well be the most difficult problem of the semester. Two students who wrongly responded "basic" to the first part of the problem subsequently responded "ph>7" on the next viewing. It is possible that these students knew the correct answer, but did not know the pH scale. Another two students repeated their mistakes in step 3 of the problem when they failed to convert K_b to K_a in both viewings. The true reason behind this mishap is unknown, but Gabel's quote may help to explain it: the feedback provided may have been attended to by the student, but not successfully committed to long-term memory because of lack of a place to store the information—lack of an existing structure of knowledge in which to place the new information. Of course other variables like cognitive load, affect, knowledge of study strategies, relevance, and self-efficacy may also have played a role. The investigation of those variables is left to future research.

Because there were so many ways to go wrong in this problem and so small a sample (with regard to those participating in and missing the problem), it is difficult to pick up trends, but one mistake that feedback could have helped prevent was the error most common across both feedback types and that had to do with the students not including hydrolysis in their reactions (1st viewing), so then not including it in their K_b expressions (2nd viewing, step 2). The hydrolysis reaction may be a key step to the problem, but it reflects an interesting set of questions in current research; Should feedback content be based on current student knowledge, cognitive load, interest, or another variable? (For sample research published after the design and implementation of this study, see Cole & Todd, 20003; Narciss, 2004; Moreno, 2004; Webber, 2004.)

The quantitative analyses revealed no significant differences between feedback types on any measure, though follow-up investigation revealed interesting patterns in participation. The results of the in-depth analyses revealed insight into the students' patterns of responses following feedback intervention. At the same time, it raised questions as to the explanation of the phenomena observed. In Chapter 5 we summarize these results and present more ideas for future work.

Chapter 5: Discussion & Conclusions

> There is something I don't know
> that I am supposed to know.
> I don't know *what* it is I don't know,
> and yet am supposed to know,
> And I feel I look stupid
> If I seem not to know it
> and not know *what* it is I don't know.
> Therefore I pretend to know everything.
> I feel you know what I am supposed to know
> but you can't tell me what it is
> because you don't know that I don't know what it is.
> --Quote from Liang, 1970, in Taylor, et al., 1984

Discussion

The quote above illustrates two of the factors that may be a part of a feedback statement: information about what one is supposed to know and the distance between what is known and what needs to be known. The poet also alludes to motivation and beliefs that may influence a lens of interpretation of the feedback, or in this case the lack thereof: one's performance-avoid motivation ("I feel I look stupid"—thus prompting the desire to know), and one's epistemological beliefs in the belief in authority dimension ("I feel you know"). These and other factors are found in the literature to play a role in student success.

Comparison of Feedback Types

This study concentrates on comparing two cognitive feedback types—informational and diagnostic—to determine which corresponds to higher course scores, higher motivation, and more mature epistemological beliefs in chemistry knowledge.

The initial presumptions led to hypotheses that diagnostic feedback might improve performance more than informational feedback because of its increased specificity and its “confront and correct” approach, which involved bringing the student’s error to attention then providing a more appropriate solution to the chemistry problem. It was also thought that diagnostic feedback, in its gentle correction of specified errors and explanation, might stimulate change in the student’s epistemological beliefs such that the recipient would be less apt to view chemistry algorithmic problems as overly simple, concepts as certain or based on a single principle, and authority as the primary source of knowledge. In a similar manner, it was thought that diagnostic feedback might promote mastery goal orientation by providing feedback that focused on the infraction rather than the student, that offered correction rather than punishment, and provided the what, why and how of the explanation. Lastly, it was thought that the beneficial aspects of diagnostic feedback would manifest most for complex, higher-level application or analysis problems since those offer the greatest opportunity for a student to become lost and frustrated or for cognitive load to be filled. The results, however, do not support any of these anticipations. Instead, we found no significant differences between feedback types on performance, motivation, or epistemological beliefs toward chemistry. We did, however, find evidence that increased participation in course assignments corresponded to higher course scores. Additionally, the degree to which a student engaged in a task, coupled with the number of e-quizzes the student participated in, related more closely to course score than the number or type of feedback treatments.

Relationships Among Feedback Specificity, Performance, Motivation, and Epistemological Beliefs in Chemistry

Regarding the objective of investigating the relationships among feedback specificity, performance, goal orientation, and epistemological beliefs, causality and path analysis could not be conducted due to small sample size (Cohen & Cohen, 1983) and the inability to transform the curvilinear functions into linear functions for use in mediator-moderator regression analyses in the Baron & Kenny (1986) style. Instead, scatter plots were created and individual nonlinear regression analyses were conducted that revealed statistically significant nonlinear relationships between course score and treatment frequency for both the diagnostic feedback group and the informational feedback group with both relationships best summarized as an inverse-U shape.

Originally, we thought the result suggested an optimal number of feedback treatments or number of questions missed that varied with feedback type. However, considering all the plots shown in Chapter 4 together, we believe that the low sample size is an issue. Also, we believe there is another variable at work, interacting with the way students receive the information and commit it to memory. The search for this variable is left for future research, though we hypothesize that it may be initial chemistry knowledge (readiness), cognitive load, or self-efficacy.

Separately, we also suspect that the relationships among the high TC recipients (Figures 4.1.1-2 and Figures 4.4.1-2) may be due, in part, to students becoming dependent on the feedback and not actively thinking for themselves, thus corresponding to lower course scores (a phenomenon known as “crutching”, see Kluger & DeNisi, 1996). But since the e-quizzes in this study were voluntary and any student could participate in any number of e-quizzes, this aspect is left for different methods in future research. Another possibility is that the students who miss many questions need assistance beyond feedback and do not possess the strategies, drive or experience to find that information or assistance for themselves (termed “information seeking” behavior). This resonates with Gabel (1999) and Kulhavy (1977) each of whom propose that only the prepared mind can fully utilize information (such as the cognitive feedback) that is given.

Looking at motivation next, we note that in both feedback groups the mastery goal orientation of students decreased significantly between the pre- and post-surveys. This finding brings questions to mind as to the reason students in this study may reduce their commitment to mastery goals. A careful look at the data revealed that A- and B-level students decreased their mastery orientation very slightly, but a small number of C-level students decreased their mastery orientation substantially. In a separate, but related, analysis we discovered that the students having low mastery orientation but high course scores, had high performance-approach orientation. It appears that students shifted slightly from goals of mastery to goals of performance-approach (concern about grades) at the end of the semester, with the C-level students appearing to have a greater shift than the A- or B-level students. (The count of D- or F-level students was too small to make any generalized conclusions.) Since the end-of-semester survey was given the week before final examinations, this result is not altogether surprising. However, we must caution interpretation of these results. While we are suggesting that motivation “shifted” from mastery to performance-approach orientation, this is just one possible interpretation. It is important to keep in mind that the goal orientations are not mutually exclusive and that it is possible to simultaneously have a high mastery and high performance-approach goal orientation. What we propose is that students have a certain amount of “energy” to allocate to tasks and that “energy” may follow conservation laws, much like chemicals do. We propose that perhaps at the end of the semester, with final examinations approaching, students allocate their energy (expressed through motivation in this study) toward making the grade rather than the “free exploration” that takes place under the mastery goals.

The performance goal orientation factor did not correlate significantly with number of corrective feedback treatments (TC) or course score for either feedback type. The performance-avoid goal orientations, representing students whose goals are to perform tasks in order to avoid embarrassment, shared a non-significant negative correlation with course score and treatment frequency for both feedback types, suggesting that higher motivation to avoid embarrassment corresponded with lower course scores and lower treatment counts (fewer misses). The relationship with course score has been detected in prior studies (Harackiewicz, 2000;VandeWalle, et al, 2001) and the relationship with treatment count makes sense when interpreted that those students who wish to avoid embarrassment and failure try not to miss quiz questions.

Epistemological beliefs were measured with a pre-post survey design for the three factors of certain/simple, authority, and attainability of truth. The post-survey measure of the simple/certain dimension showed a significant positive correlation with TC (direct proportion) for the informational feedback group. A non-significant negative (inverse) correlation was found between the certain/simple dimension and TC for the diagnostic feedback type.

One epistemological beliefs variable, the belief in the attainability of truth, showed a correlation with e-quiz scores. The naïve view believes that absolute truth is attainable while the mature view has a more “relative” belief about truth, to the point that truth can be formed in oneself. Initial belief in the attainability of truth showed a negative correlation between its pre-survey measure and e-quiz scores, such that more naïve beliefs correlated with lower course scores (Table 4.6). That information, coupled with the results showing that students’ initial beliefs that absolute truth exists impacted their scores for online assignments (e-quizzes) but not their course grade that was derived from in-class examinations, leads us to wonder whether students' beliefs in the attainability of truth—and behavior in response to those beliefs—differ between face-to-face instruction and web-based instruction. This new measure may have implications for distance learning. However, caution is warranted since these results are based on only two survey questions and these results have not been replicated.

Also, though not shown, we discovered that initial belief in the attainability of truth (mTRU scores) interacted with mastery goal orientation, which suggests that the observed change in mastery orientation (MGO) was not consistent across mTRU levels. It appears that students believing strongly that truth is attainable showed a change in MGO that was different than those who didn't share such a belief.

Like the statistical analyses, the in-depth per-individual analyses of case studies showed a trend of similar performance between the two feedback types, although informational feedback appears to be the stronger choice in Case 3, involving the very difficult pH of a salt problem. The data for that case seem to indicate that informational feedback is associated with increased performance on examinations for problems that were different, but related, to the practice problems. While one must be wary in taking the results of so few case studies too far, they did present interesting findings in the successive answers students gave in practice. For one, while general rates of success didn't change much: students appeared to pay attention to the feedback in that they did not tend to respond with the same wrong answer twice. Thus the lack of change in overall success rates on subsequent problems appears not to be due to students replicating mistakes, but rather an incremental addition of understanding. It appears that the two feedback types helped the students to the next step in the sequence to solve the problem, which suggests that feedback may be playing a Vygotskian role (1978; see also Fakhreddine, 1999). Vygotsky (1978) proposed that students learn best when teamed up with someone possessing incrementally more knowledge in a given area. The maximum incremental "distance" was defined as the Zone of Proximal Development, or ZPD (Vygotsky, 1978). Matches of mentor-student in which the distance was outside the student's ZPD proved less fruitful than those within it. Now, applying that to feedback, it would be interesting to know whether the feedback provided in this study included some information that was within the student's ZPD and much of it outside, so that the student absorbed only the next step of the sequence and could not absorb the rest of the information provided. This represents one possible explanation of the observations noted in the case studies in Chapter 4.

Limitations

Diagnostic feedback, as used in this study, employed a “confront and correct” strategy for errors. The assumed error was challenged and then a correct alternative solution was given. There were several problems in implementing this approach. The questions used in this study originated from a chemistry database. The questions included those written to accompany a popular chemistry textbook and questions added periodically by the instructors at the university.

The challenge came in writing the per-response feedback. The database in use has collected, over time, the most popular incorrect student responses and uses those as distracters. In order to determine how a student might have arrived at such a wrong answer (indeed, necessary for diagnosis of the error), the problem had to be worked backwards from each response. Sometimes the error could be determined and other times it could not. Only popular, known, or systematic student errors were included as distracters for this study. This limits the research greatly since, as any teacher can attest to, students do not always misunderstand a concept or go astray on calculations in predictable ways. Another limitation may arise from the wording of the feedback. In writing the feedback, we sometimes inadvertently referred to the student by saying, for example, “It looks like you multiplied mass times volume…”. As described previously, effective feedback (Ilgen, et al, 1979) focuses on correcting the *task* rather than the person, such that it is better to restrain the use of the pronoun “you”. This error was not committed consistently throughout the study. An example of this change of focus is located in Appendix A.3, comparing feedback provided for Case Study 2 and Case Study 3.

Further limitation lies in the fact that we did not distinguish or capture the reason behind a student’s response. That is, we did not distinguish between students having an ontological impediment from missing or disconnected knowledge and those who guessed. We distinguished only between those who missed the problem and those who did not. The reasons behind a student missing a problem can be many and we did not distinguish among those reasons.

Future Work

Besides those recommendations already mentioned, we propose three additional items for future research:

(1) the creation of a database of problems, including a collection of student work and responses, to serve as a clearinghouse of chemical miscalculations and misconceptions to be shared among researchers and instructors alike,

(2) a test of feedback specificity using another type of cognitive feedback—goal-oriented feedback ("realign and redefine")—that includes statements to help the students to orient their studies to meet the objectives of the assignment, in lieu of diagnostic feedback,

(3) a pre-post study design for performance measures, to help distinguish between "willing" and "capable" acceptance of feedback,

(4) determination of students' preferred feedback type(s), and

(5) distinguish among students' responses by error committed

We propose that a database of problems, including student errors and student solutions arriving at those errors, would help greatly in future feedback (and assessment) research. We recognize that the thousands of students taking freshman chemistry each year at our institution are far more creative in arriving at erroneous results than we were at mimicking them.

With regard to goal-oriented cognitive feedback, a limitation is that assignments will need to be more focused in nature (covering few topics) than were afforded in this field study. Problems will need to be more carefully chosen and matched to specific learning objectives than is found in a typical college classroom setting. At present, a majority of the instructors appear to present a large quantity of material in any given lecture, then present quizzes or homework assignments weekly or biweekly. This results in a great many topics included in any given assignment, with the apparent purpose to expose the students to a range of problems on a given set of topics rather than restricting problems to specific learning objectives. The question for chemical education research is whether the addition of goals to the assignment and the feedback will increase student performance on the assignment and overall. Such feedback has been recommended and utilized for learning in other contexts (Locke and Latham, 1990; Butler and Nisan, 1986). Since our use of diagnostic feedback may have been viewed as controlling, a focus on goals may “skirt” that issue.

The third recommendation involves student readiness versus willingness. The acceptability of feedback relies somewhat on the perceived value and positive nature of it. Would a well-prepared student consider corrective feedback for a given problem more positively than an under-prepared student receiving the same feedback on the same problem?

The fourth recommendation focuses on the students, to determine which level of feedback they prefer to receive and whether the chosen level of feedback varies with question difficulty, motivation, background knowledge and proficiency, content, self-efficacy, or metacognition (or some combination of these).

The fifth and final recommendation responds to the last of the limitations discussed in this chapter: the fact that we did not distinguish among errors made based on misconceptions, miscalculations, and guesses. Since the effect of feedback on performance could depend on the information value of the feedback, future researchers may choose to distinguish between “guessers” and “ontological impediments”, for example, to get a more focused view of the effect of feedback on performance.

Summary

This study focuses on the comparison of diagnostic (per response) and informational (per question) feedback and found that diagnosis of the error may not be needed if the informational feedback is informative enough. The quasi-qualitative results suggest that some students were not prepared enough to internalize the feedback provided, since students who answered in one way on the first viewing answered with understanding only of one more step of the solution on the next viewing. It was suggested that such behavior might reflect either ill-formed knowledge structures or cognitive load limitations. Taking a step back, these results also suggest that teachers can utilize planned distracters to help students self-assess their knowledge, an important aspect of effective feedback.

It appears that the added cost and effort to provide diagnostic feedback is not warranted when compared with the results for informational feedback. Students receiving diagnostic feedback did not outperform those receiving informational feedback, regardless of whether performance was analyzed on a question-feedback-question scale (case studies), on a difficulty of question scale, on e-quiz scores per unit scale, or as manifested in course score. Even so, students in the diagnostic feedback group who made predictable, systematic errors in calculations received feedback attuned to their error. While the diagnostic feedback did not correspond to increased performance, motivation, or epistemological beliefs, it also did not correspond to reduced scores for those measures.

Appendices

Appendix A.1

Achievement Goal Theory Survey

and

Epistemological Beliefs Questionnaire

Appendix A.1.1: Achievement Goal Orientation Survey

1: Strongly Disagree; 2: Disagree; 3: Neutral; 4: Agree; and 5: Strongly Agree

These statements should be taken as straightforward and simple descriptions of your attitudes. If you think the statement is very true of you, bubble 5; if a statement is not at all true of you, bubble 1. If the statement is more or less true of you, find the number between 1 and 5 that best describes you.

1	The reason I will do the chemistry quizzes is so the instructor won't think I know less than others.	1	2	3	4	5
2	I will feel successful in the chemistry class if I do better than most of the other students in the class.	1	2	3	4	5
3	I will do the chemistry assignments because I am interested in them.	1	2	3	4	5
4	I will do the chemistry assignments so others in the class won't think I'm ignorant.	1	2	3	4	5
5	I want to do better than other students in my chemistry class.	1	2	3	4	5
6	One reason I might not participate in the chemistry quizzes is to avoid looking incompetent.	1	2	3	4	5
7	Doing better than other students in the chemistry class is important to me.	1	2	3	4	5
8	One of my main goals in the che4mistry quizzes will be to avoid looking like I can't do my work.	1	2	3	4	5
9	I'd like to show my chemistry instructor that I'm smarter than the other students in the physical science class.	1	2	3	4	5
10	It's very important to me that I don't look incompetent in my chemistry class.	1	2	3	4	5

11	I like working on more difficult quiz questions that I'll learn more from, even if I make a lot of mistakes.	1	2	3	4	5
12	An important reason I do my chemistry work is so I won't embarrass myself.	1	2	3	4	5
13	An important reason I participate in the chemistry assignments is because I want a deeper understanding of chemistry.	1	2	3	4	5
14	I will feel really good if I will be the only one who will be able to answer the assignment questions correctly.	1	2	3	4	5
15	I like assignments best with they really make me think.	1	2	3	4	5
16	Truth is unchanging in this subject.	1	2	3	4	5

Appendix A.1.2: Epistemological Beliefs Questionnaire

The following 24 statements may or may not describe your feelings and beliefs ***about knowing and knowing chemistry.*** Please rate each statement by bubbling in the number between 1 and 5 according to the following scale:

1: Strongly Disagree; 2: Disagree; 3: Neutral; 4: Agree; and 5: Strongly Agree

These statements should be taken as straightforward and simple descriptions of your beliefs. If you think the statement is very true of you, bubble 5; if a statement is not at all true of you, bubble 1. If the statement is more or less true of you, find the number between 1 and 5 that best describes you.

17	In this subject, most work has only one right answer.	1	2	3	4	5
18	Sometimes you just have to accept answers from the experts in this field, even if you don't understand them.	1	2	3	4	5
19	All professors in this field would probably come up with the same answers to questions in this field.	1	2	3	4	5
20	If you read something in a textbook for this subject, you can be sure it is true.	1	2	3	4	5
21	Most of what is true in this subject is already known.	1	2	3	4	5
22	In this subject, it is good to question the ideas presented.	1	2	3	4	5
23	If scholars try hard enough, they can find the answers to almost anything.	1	2	3	4	5
24	Experts in this field can ultimately get to the truth.	1	2	3	4	5
25	Principles in this field are unchanging.	1	2	3	4	5
26	If my personal experience conflicts with ideas in the textbook, the book is probably right.	1	2	3	4	5
27	Answers to questions in this field change as experts gather more information.	1	2	3	4	5
28	All experts in this field understand the field in the same way.	1	2	3	4	5
29	I am most confident that I know something when I know what the experts think.	1	2	3	4	5

Appendix A.2

Factor Assignments for Survey Questions

Variable	Motivation				Epistemological beliefs		
Instrument	Milner-Bolotin Survey Questions				Hofer Survey Questions		
Factors	MGO	SHOW	Av1	Av2	Cert	Auth	Tru
Question Numbers	3, 11, 13, 15	2, 5, 7, 9, 14	6, 8, 10, 12	1, 4	16, 17, 19, 21, 22*, 25, 27*, 28	18, 20, 26, 29	23, 24

* Denotes questions that were reverse-coded before averaging

Appendix A.3

Examples of Diagnostic and Informational Feedback

Case Study 1:
Conceptual Buffer Problem (All 3 viewings)

A buffer is composed of NH_3 and NH_4Cl. How would this buffer solution control the pH of a solution when a small amount of a strong base is added?

A. The OH^- reacts with the NH_4^+.
B. The OH^- reacts with the Cl^-.
C. The OH^- reacts with the NH_3.
D. The OH^- reacts with the H_2O.

Informational Feedback

Informational Feedback for Response A:
Correct. Cl^- is a spectator ion and does not react. The NH_3 is the base, B, and NH_4^+ is the conjugate acid BH^+. So when base is added, it will react with the acid.
Informational Feedback for Responses B, C, D:
Incorrect. Cl^- is a spectator ion and does not react. The NH_3 is the base, B, and NH_4^+ is the conjugate acid BH^+. So when base is added, it will react with the acid.

Diagnostic Feedback

Diagnostic Feedback for Response A:
Correct. Cl^- is a spectator ion and does not react. The NH_3 is the base, B, and NH_4^+ is the conjugate acid BH^+. So when base is added, it will react with the acid.

Diagnostic Feedback for Response B:
Incorrect. How can the Cl^-, an anion, react with OH^-, another anion? Perhaps you misread the statement. Since like charges repel, there is no driving force for two anions to combine.
Cl^- is a spectator ion and does not react. The NH_3 is the base, B, and NH_4^+ is the conjugate acid BH^+. So when base is added, it will react with the acid.

Diagnostic Feedback for Response C:
Incorrect. As a weak base, NH_3 is willing to accept a proton, but is not willing to accept a hydroxide ion. So that would not happen. There is, however, a cation in the solution that the added OH^- could react with!
Cl^- is a spectator ion and does not react. The NH_3 is the base, B, and NH_4^+ is the conjugate acid BH^+. So when base is added, it will react with the acid.

Diagnostic Feedback for Response D:
Incorrect. Perhaps you were thinking of the reaction
$2H_2O \leftarrow\rightarrow H_3O^+ + OH^-$
Notice that in that reaction OH^- is reacting with hydronium, not water. There is no effective way for a water molecule to accept (react with) OH^-.

Cl^- is a spectator ion and does not react. The NH_3 is the base, B, and NH_4^+ is the conjugate acid BH^+. So when base is added, it will react with the acid.

Case Study 2:
Reaction Rate (1st and 2nd Viewings; on e-quiz)

Consider the multistep reaction that has the overall reaction
2A + 2B --> C + D.
What is the rate law expression that would correspond to the following proposed mechanism?

A + B --> I (fast)
I + B --> C + X (slow)
A. X + A --> D (fast)

B. Rate = $k[A][B]^2$
C. Rate = k[I][B]
D. Rate = k[A][B][I]
E. Rate = k[A][B][I][X][D]

Informational Feedback

Informational Feedback for Response A:
Correct! 'The experimentally determined reaction orders of reactants indicate the number of molecules of those reactants involved in

- the slow step only, if it occurs first, or
- the slow step and any fast equilibrium steps preceding the slow step.'

Writing rate eqn for slow step above gives:
Rate = k[I][B]
The preceding step is a fast equilibrium step, so
Rate forward = kf[A][B] = kr[I] = rate reverse
so then [I] = kf/kr * [A][B]
can be substituted into the rate equation above to give the rate law.

Informational Feedback for Responses B, C, D:
Incorrect. 'The experimentally determined reaction orders of reactants indicate the number of molecules of those reactants involved in

- the slow step only, if it occurs first, or
- the slow step and any fast equilibrium steps preceding the slow step.'

Writing rate eqn for slow step above gives:
Rate = k[I][B]
The preceding step is a fast equilibrium step, so
Rate forward = kf[A][B] = kr[I] = rate reverse
so then [I] = kf/kr * [A][B]
can be substituted into the rate equation above to give the rate law.

Diagnostic Feedback

Diagnostic Feedback for Response A:
Correct! 'The experimentally determined reaction orders of reactants indicate the number of molecules of those reactants involved in

- the slow step only, if it occurs first, or

- the slow step and any fast equilibrium steps preceding the slow step.'

Writing rate eqn for slow step above gives:
Rate = k[I][B]
The preceding step is a fast equilibrium step, so
Rate forward = kf[A][B] = kr[I] = rate reverse
so then [I] = kf/kr * [A][B]
can be substituted into the rate equation above to give the rate law.

Diagnostic Feedback for Response B:
Incorrect. Yes, it appears you have correctly identified the slow step of the reaction mechanism. Good work. However, the reaction rate must be written in terms of the reactants. And it must include the slow step and any preceding fast equilibrium steps. Like most cases, there are several ways to do this problem. Here is one:
'The experimentally determined reaction orders of reactants indicate the number of molecules of those reactants involved in

- the slow step only, if it occurs first, or
- the slow step and any fast equilibrium steps preceding the slow step.'

Writing rate eqn for slow step above gives:
Rate = k[I][B]
The preceding step is a fast equilibrium step, so
Rate forward = kf[A][B] = kr[I] = rate reverse
so then [I] = kf/kr * [A][B]
can be substituted into the rate equation above to give the rate law.

Diagnostic Feedback for Response C:
Nice try, but not quite. The [I] cancels out.
'The experimentally determined reaction orders of reactants indicate the number of molecules of those reactants involved in

- the slow step only, if it occurs first, or
- the slow step and any fast equilibrium steps preceding the slow step.'

Writing rate eqn for slow step above gives:
Rate = k[I][B]
The preceding step is a fast equilibrium step, so
Rate forward = kf[A][B] = kr[I] = rate reverse
so then [I] = kf/kr * [A][B]
can be substituted into the rate equation above to give the rate law.

Diagnostic Feedback for Response D:
Incorrect. While position of equilibrium can vary with concentrations of substances, the rate does not have to. The rate equation does not have to include all of the substances in a mechanism.
'The experimentally determined reaction orders of reactants indicate the number of molecules of those reactants involved in

- the slow step only, if it occurs first, or

- the slow step and any fast equilibrium steps preceding the slow step.'

Writing rate eqn for slow step above gives:
Rate = k[I][B]
The preceding step is a fast equilibrium step, so
Rate forward = kf[A][B] = kr[I] = rate reverse
so then [I] = kf/kr * [A][B]
can be substituted into the rate equation above to give the rate law.

Case Study 3: pH of a Salt
1st viewing; Step 1

A question states:
What is the pH of a 0.12 M solution of anilinium nitrate, $C_6H_5NH_3NO_3$? K_b for aniline ($C_6H_5NH_2$) is 4.2×10^{-10}.
Will the solution be acidic, basic, or neutral? Or is it predictable at all?

A. acidic
B. basic
C. neutral
D. cannot be predicted from the information given

Informational Feedback

Informational Feedback for Response A:
Correct! It is a salt whose "parentage" is a strong acid and a weak base, so it will produce an acidic solution. Thus, any pH we ultimately choose had better be acidic (pH<7).

Informational Feedback for Responses B, C, or D:
Incorrect. It is a salt whose "parentage" is a strong acid and a weak base, so it will produce an acidic solution. Thus, any pH we ultimately choose had better be acidic (pH<7).

Diagnostic Feedback

Diagnostic Feedback for Response A:
Correct! It is a salt whose "parentage" is a strong acid and a weak base, so it will produce an acidic solution. Thus, any pH we ultimately choose had better be acidic (pH<7).

Diagnostic Feedback for Response B:
Incorrect. Perhaps you forgot that HNO_3 is a strong acid? Or perhaps you're still not sure how to tell?
Well, take the salt and split it into its cation and its anion.
To the anion, add H^+. That's the acid parent.
To the cation, add OH^- or subtract an H^+. (In this case, do that latter. How do I know? Because I'm given info on aniline.) That's the basic parent.
The acid parent is a strong acid. The basic parent is a weak base (it has a K_b value).
Since it is a salt whose "parentage" is a strong acid and a weak base, it will produce an acidic solution. Thus, any pH we ultimately choose had better be acidic (pH<7).

Diagnostic Feedback for Response C:
Incorrect. Only salts formed from the reaction of a strong acid and a strong base can form neutral solutions.
Or perhaps you're still not sure how to tell?
Well, take the salt and split it into its cation and its anion.
To the anion, add H^+. That's the acid parent.

To the cation, add OH^- or subtract an H^+. (In this case, do that latter. How do I know? Because I'm given info on aniline.) That's the basic parent.
The acid parent is a strong acid. The basic parent is a weak base (it has a K_b value).
Since it is a salt whose "parentage" is a strong acid and a weak base, it will produce an acidic solution. Thus, any pH we ultimately choose had better be acidic (pH<7).

Diagnostic Feedback for Response D:
Actually, we can! Just look at the salt's "parentage". Oh, not sure how to tell?
Well, take the salt and split it into its cation and its anion.
To the anion, add H^+. That's the acid parent.
To the cation, add OH^- or subtract an H^+. (In this case, do that latter. How do I know? Because I'm given info on aniline.) That's the basic parent.
The acid parent is a strong acid. The basic parent is a weak base (it has a K_b value).
Since it is a salt whose "parentage" is a strong acid and a weak base, it will produce an acidic solution. Thus, any pH we ultimately choose had better be acidic (pH<7).

Case Study 3
First viewing, Step 2

Continuing the question:
What is the pH of a 0.12 M solution of anilinium nitrate, $C_6H_5NH_3NO_3$? K_b for aniline ($C_6H_5NH_2$) is 4.2×10^{-10}.
It is apparent that we must determine $[H_3O^+]$ in order to arrive at pH (that's one way to tackle this problem). Which reaction will lead us to that concentration?

A. $C_6H_5NH_3^+ + H_2O \leftarrow\rightarrow C_6H_5NH_2 + H_3O^+$
B. $HNO_3 + H_2O \leftarrow\rightarrow NO_3^- + H_3O^+$
C. $C_6H_5NH_3NO_3 \leftarrow\rightarrow C_6H_5NH_3^+ + NO_3^-$
D. $2\ H_2O \leftarrow\rightarrow H_3O^+ + OH^-$
E. $C_6H_5NH_3NO_3 + H_2O \leftarrow\rightarrow C_6H_5NH_3NO_3 + H_3O^+$

Informational Feedback

Informational Feedback for Response A:
Correct. We've already determined that the given salt, anilinium nitrate, is said to have a weak base and a strong acid as "parents". When the salt dissociates in water and then reacts with water, two reactions are possible:

- $NO_3^- + H_2O$ <- -> $HNO_3 + OH^-$

and

- $C_6H_5NH_3^+ + H_2O$ <- -> $C_6H_5NH_2 + H_3O^+$

The first reaction goes only in the reverse…that is, recall that HNO_3 is a strong acid, so it is nearly completely dissociated. So the base OH^- is not produced to any appreciable extent.
The second reaction, though, can occur because the aniline formed is a weak base (we are given that info).
Also note that the forward reaction gives us H_3O^+, which will cause a change in pH that we can soon calculate.

Informational Feedback for Responses B, C, D, or E:
Incorrect. We've already determined that the given salt, anilinium nitrate, is said to have a weak base and a strong acid as "parents". When the salt dissociates in water and then reacts with water, two reactions are possible:

- $NO_3^- + H_2O$ <- -> $HNO_3 + OH^-$

and

- $C_6H_5NH_3^+ + H_2O$ <- -> $C_6H_5NH_2 + H_3O^+$

The first reaction goes only in the reverse…that is, recall that HNO_3 is a strong acid, so it is nearly completely dissociated. So the base OH^- is not produced to any appreciable extent.
The second reaction, though, can occur because the aniline formed is a weak base (we are given that info).
Also note that the forward reaction gives us H_3O^+, which will cause a change in pH that we can soon calculate.

Diagnostic Feedback

Diagnostic Feedback for Response A:

Correct. We've already determined that the given salt, anilinium nitrate, is said to have a weak base and a strong acid as "parents". When the salt dissociates in water and then reacts with water, two reactions are possible:

- NO_3^- + H_2O <- -> HNO_3 + OH^-

and

- $C_6H_5NH_3^+$ + H_2O <- -> $C_6H_5NH_2$ + H_3O^+

The first reaction goes only in the reverse…that is, recall that HNO_3 is a strong acid, so it is nearly completely dissociated. So the base OH^- is not produced to any appreciable extent.

The second reaction, though, can occur because the aniline formed is a weak base (we are given that info).

Also note that the forward reaction gives us H_3O^+, which will cause a change in pH that we can soon calculate.

Diagnostic Feedback for Response B:

Incorrect. While you are correct that this reaction has the hydronium ion product, it is not the equation we seek. Why not?

Let's look at a mathematical argument:

If we were to use that chemical equation, we would see that

$K_a = [H_3O^+][NO_3^-] / [HNO_3]$

We'd need the K_a so that we can solve for x. But there is no K_a listed for nitric acid. Why not? Because the K_a is "very large". In other words, the forward reaction goes very nearly to completion, so much so that the reverse reaction happens very little in comparison.

So, this chemical equation is not the one we seek.

Here's the better way to think about this question:

We've already determined that the given salt, anilinium nitrate, is said to have a weak base and a strong acid as "parents". When the salt dissociates in water and then reacts with water, two reactions are possible:

- NO_3^- + H_2O <- -> HNO_3 + OH^-

and

- $C_6H_5NH_3^+$ + H_2O <- -> $C_6H_5NH_2$ + H_3O^+

The first reaction goes only in the reverse…that is, recall that HNO_3 is a strong acid, so it is nearly completely dissociated. So the base OH^- is not produced to any appreciable extent.

The second reaction, though, can occur because the aniline formed is a weak base (we are given that info).

Also note that the forward reaction gives us H_3O^+, which will cause a change in pH that we can soon calculate.

Diagnostic Feedback for Response C:

Incorrect. You have correctly chosen the dissociation equation. However, the question asked for the chemical equation that determines the pH of the solution. See if this helps you to think about this question:

We've already determined that the given salt, anilinium nitrate, is said to have a weak base and a strong acid as "parents". When the salt dissociates in water and then reacts with water, two reactions are possible:

- $NO_3^- + H_2O$ <- -> $HNO_3 + OH^-$

and

- $C_6H_5NH_3^+ + H_2O$ <- -> $C_6H_5NH_2 + H_3O^+$

The first reaction goes only in the reverse…that is, recall that HNO_3 is a strong acid, so it is nearly completely dissociated. So the base OH^- is not produced to any appreciable extent.
The second reaction, though, can occur because the aniline formed is a weak base (we are given that info).
Also note that the forward reaction gives us H_3O^+, which will cause a change in pH that we can soon calculate.

Diagnostic Feedback for Response D:
Incorrect. This is the equation for water and does not help us determine the pH of this solution since there is a larger contributor to pH (the salt). See if this explanation helps:
We've already determined that the given salt, anilinium nitrate, is said to have a weak base and a strong acid as "parents". When the salt dissociates in water and then reacts with water, two reactions are possible:

- $NO_3^- + H_2O$ <- -> $HNO_3 + OH^-$

and

- $C_6H_5NH_3^+ + H_2O$ <- -> $C_6H_5NH_2 + H_3O^+$

The first reaction goes only in the reverse…that is, recall that HNO_3 is a strong acid, so it is nearly completely dissociated. So the base OH^- is not produced to any appreciable extent.
The second reaction, though, can occur because the aniline formed is a weak base (we are given that info).
Also note that the forward reaction gives us H_3O^+, which will cause a change in pH that we can soon calculate.

Diagnostic Feedback for Response E:
Incorrect. If the given substance were a weak acid, then you would be correct. But the given substance is a salt. In an earlier problem we determined that the given salt, anilinium nitrate, is said to have a weak base and a strong acid as "parents".
When the salt dissociates in water and then reacts with water, two reactions are possible:

- $NO_3^- + H_2O$ <- -> $HNO_3 + OH^-$

and

- $C_6H_5NH_3^+ + H_2O$ <- -> $C_6H_5NH_2 + H_3O^+$

The first reaction goes only in the reverse…that is, recall that HNO_3 is a strong acid, so it is nearly completely dissociated. So the base OH^- is not produced to any appreciable extent.

The second reaction, though, can occur because the aniline formed is a weak base (we are given that info).
Also note that the forward reaction gives us H_3O^+, which will cause a change in pH that we can soon calculate.

Case Study 3
First viewing, Step 3

Now let's finish out the question.
What is the pH of a 0.12 M solution of anilinium nitrate, $C_6H_5NH_3NO_3$?
K_b for aniline ($C_6H_5NH_2$) is 4.2 x 10^{-10}.

A. 2.77
B. 5.15
C. 11.23
D. None of these
E. Unable to determine.

Informational Feedback

Informational Feedback for Response A:
Correct! $K_a = K_w/K_b = [C_6H_5NH_2][H_3O^+] / [C_6H_5NH_3NO_3]$
$K_a = (1 \times 10^{-14})/(4.2 \times 10^{-10}) = x^2 / (0.12-x)$
assume that 0.12-x = 0.12
x = 1.69 x 10^{-3}
Then calculate pH from that! (It is better to not round off values until the very end of the calculation.)

Informational Feedback for Responses B, C, D, or E:
Incorrect. $K_a = K_w/K_b = [C_6H_5NH_2][H_3O^+] / [C_6H_5NH_3NO_3]$
$K_a = (1 \times 10^{-14})/(4.2 \times 10^{-10}) = x^2 / (0.12-x)$
assume that 0.12-x = 0.12
x = 1.69 x 10^{-3}
Then calculate pH from that! (It is better to not round off values until the very end of the calculation.)

Diagnostic Feedback

Diagnostic Feedback for Response A:
Correct! $K_a = K_w/K_b = [C_6H_5NH_2][H_3O^+] / [C_6H_5NH_3NO_3]$
$K_a = (1 \times 10^{-14})/(4.2 \times 10^{-10}) = x^2 / (0.12-x)$
assume that 0.12-x = 0.12
x = 1.69 x 10^{-3}
Then calculate pH from that! (It is better to not round off values until the very end of the calculation.)

Diagnostic Feedback for Response B:
Incorrect. Please check your work: did you convert the given K_b to K_a? The chemical equation determined in the previous question was a K_a equation (can tell because acidic product and conjugate base reacting with water as reactants). But we're given the value of K_b. So we need to convert K_b to K_a in order to proceed with the calculation.
$K_a = K_w/K_b = [C_6H_5NH_2][H_3O^+] / [C_6H_5NH_3NO_3]$
$K_a = (1 \times 10^{-14})/(4.2 \times 10^{-10}) = x^2 / (0.12-x)$
assume that 0.12-x = 0.12
x = 1.69 x 10^{-3}

Then calculate pH from that! (It is better to not round off values until the very end of the calculation.)

Diagnostic Feedback for Response C:

Incorrect. This cannot be the answer because we decided in the first steps that the given substance, anilinium nitrate, is a salt of a weak base/strong acid, so it would create an acidic solution when added to water. This pH value is basic, so cannot be the answer.

$K_a = K_w/K_b = [C_6H_5NH_2][H_3O^+] / [C_6H_5NH_3NO_3]$

$K_a = (1 \times 10^{-14})/(4.2 \times 10^{-10}) = x^2 / (0.12-x)$

assume that 0.12-x = 0.12

$x = 1.69 \times 10^{-3}$

Then calculate pH from that! (It is better to not round off values until the very end of the calculation.)

Diagnostic Feedback for Response D:

Actually, the answer is listed. Please check your work against this:

$K_a = K_w/K_b = [C_6H_5NH_2][H_3O^+] / [C_6H_5NH_3NO_3]$

$K_a = (1 \times 10^{-14})/(4.2 \times 10^{-10}) = x^2 / (0.12-x)$

assume that 0.12-x = 0.12

$x = 1.69 \times 10^{-3}$

Then calculate pH from that! (It is better to not round off values until the very end of the calculation.)

Case Study 3: pH of a Salt
Second viewing; Step 1

A question asks:
What is the pH of a 0.12 M solution of ammonium nitrate, NH_4NO_3. K_b for ammonia, NH_3, is 1.8×10^{-5}.
Will the pH be acidic, basic, neutral, or cannot be determined?

A. pH < 7
B. pH > 7
C. pH = 7
D. cannot be determined from the information given

Informational Feedback

Informational Feedback for Response A:
Correct. NH_4NO_3 is a salt. Its “parentage” is NH_4OH and HNO_3, a weak base and a strong acid. Such salts form acidic solutions.

Informational Feedback for Response B:
Incorrect. NH_4NO_3 is a salt. Its “parentage” is NH_4OH and HNO_3, a weak base and a strong acid. Such salts form acidic solutions.

Diagnostic Feedback

Diagnostic Feedback for Response A:
Correct. NH_4NO_3 is a salt. Its “parentage” is NH_4OH and HNO_3, a weak base and a strong acid. Such salts form acidic solutions.

Diagnostic Feedback for Response B:
Incorrect. The solution is not basic. It becomes acidic when the salt is added to it. I can predict that by analyzing the “parentage” of the salt: NH_4OH and HNO_3, a weak base and a strong acid. Such salts form acidic solutions.

Diagnostic Feedback for Response C:
Incorrect. If the salt had strong acid/strong base “parents” then yes, the solution would remain neutral. But this salt does not. NH_4OH and HNO_3, a weak base and a strong acid. Such salts form acidic solutions.

Diagnostic Feedback for Response D:
Actually, we can predict general pH by analyzing the reactions a salt undergoes in solution, or looking at a “shortcut”
NH_4NO_3 is a salt. Its “parentage” is NH_4OH and HNO_3, a weak base and a strong acid. Such salts form acidic solutions.

Case Study 3: pH of a Salt
Second viewing; Step 2

A question asks:
What is the pH of a 0.12 M solution of ammonium nitrate, NH_4NO_3. K_b for ammonia, NH_3, is 1.8 x 10^{-5}.
What is the K expression to use to calculate $[H_3O^+]$ directly? (Recall that brackets [ ] denote concentration.)

A. $K_a = [NH_3][H_3O^+] / [NH4^+]$
B. $K_b = [HNO_3][OH^-] / [NO_3^-]$
C. $K_b = [NH_4^+][OH^-] / [NH_3]$
D. $K_b = [NH_3][HNO_3] / [NH_4NO_3]$
E. None of these

Informational Feedback

Informational Feedback for Response A:
Good thinking. K_a for ammonium is based on the reaction NH_4^+ + H_2O <- -> H_3O^+ + NH_3
And recall that we can calculate K_a from the given K_b because ammonium and ammonia are a conjugate acid/base pair. $K_a = K_w/K_b$.

Informational Feedback for Responses B, C, D, or E:
Incorrect. K_a for ammonium is based on the reaction NH_4^+ + H_2O <- -> H_3O^+ + NH_3
And recall that we can calculate K_a from the given K_b because ammonium and ammonia are a conjugate acid/base pair. $K_a = K_w/K_b$.

Diagnostic Feedback

Diagnostic Feedback for Response A:
Good thinking. K_a for ammonium is based on the reaction NH_4^+ + H_2O <- -> H_3O^+ + NH_3
And recall that we can calculate K_a from the given K_b because ammonium and ammonia are a conjugate acid/base pair. $K_a = K_w/K_b$.

Diagnostic Feedback for Response B:
Incorrect. This equation cannot calculate $[H_3O^+]$ at all. K_b for nitric acid is "very large" and the $[OH^-]$ that would be in the equation would be so small as to give a false pH.
See if this explanation helps:
K_a for ammonium is based on the reaction
NH_4^+ + H_2O <- -> H_3O^+ + NH_3
And recall that we can calculate K_a from the given K_b because ammonium and ammonia are a conjugate acid/base pair. $K_a = K_w/K_b$.

Diagnostic Feedback for Response C:
Incorrect. You've written the K_b expression perfectly, but it is not the best answer to the question of what K expression will lead us to hydronium ion concentration **directly**. Here is one explanation of the K_a expression:
K_a for ammonium is based on the reaction

$NH_4^+ + H_2O$ <- -> $H_3O^+ + NH_3$
And recall that we can calculate K_a from the given K_b because ammonium and ammonia are a conjugate acid/base pair. $K_a = K_w/K_b$.

Diagnostic Feedback for Response D:
Incorrect. Ammonium nitrate cannot be the reactant in a K_a or K_b expression because it is not an acid or a base: it is a salt. (This is only one reason why that cannot be the correct expression.)
See if this explanation helps:
K_a for ammonium is based on the reaction
$NH_4^+ + H_2O$ <- -> $H_3O^+ + NH_3$
And recall that we can calculate K_a from the given K_b because ammonium and ammonia are a conjugate acid/base pair. $K_a = K_w/K_b$.

Diagnostic Feedback for Response E:
Actually, the correct expression is listed. Keep in mind that we can calculate K_a from the given K_b because ammonium and ammonia are a conjugate acid/base pair. $K_a = K_w/K_b$.
And when in doubt, you can take the salt, break it up into its cation and anion parts, and react each part with water. One reaction will give you an OH^- product. That's the reaction for K_b using that salt. The other reaction will give you an H_3O^+ product. It's the K_a reaction for that salt.
Next, analyze the conjugate acid for the anion part. Is it a strong acid? In our case, YES. Since it is strong, the reverse reaction happens (so no extra OH^- from that reaction.)
Next, analyze the conjugate base for the cation part. Is it a strong base? In our case, NO. (I know because I'm given a K_b for NH_3.) Ah, so the forward and reverse reactions BOTH occur and there is extra H_3O^+ in solution due to that reaction.
So, that's the relevant reaction: the K_a reaction for that salt—because that's the reaction that causes a change in pH. Here's a shorter summary:
K_a for ammonium is based on the reaction
$NH_4^+ + H_2O$ <- -> $H_3O^+ + NH_3$
And recall that we can calculate K_a from the given K_b because ammonium and ammonia are a conjugate acid/base pair. $K_a = K_w/K_b$.

Case Study 3: pH of a Salt
Second viewing; Step 3

A question asks:
What is the pH of a 0.12 M solution of ammonium nitrate, NH_4NO_3. K_b for ammonia, NH_3, is 1.8 x 10^{-5}.
Calculate the pH of the solution.

A. 5.09
B. 2.83
C. 7.00
D. 11.2

Informational Feedback

Informational Feedback for Response A:
Yes, indeed! Ka for ammonium is based on the reaction
$NH_4^+ + H_2O$ <- -> $H_3O^+ + NH_3$
Now we have H^+.
We can calculate K_a from K_b.
$K_w = K_a * K_b$
$K_a = K_w/K_b = 1 \times 10^{-14} / 1.8 \times 10^{-5} = 5.555556 \times 10^{-10}$
So that
$Ka = 5.555556 \times 10^{-10} = [H_3O^+][NH_3] / [NH_4^+]$
and $5.555556 \times 10^{-10} = x^2 / (0.12-x)$
but we can approximate that 0.12-x is about equal to 0.12
(and then verify our assumption at the end)
$x = 8.16497 \times 10^{-6}$
pH = $-(\log [H^+]) = 5.09$

Informational Feedback for Responses B, C, or D:
Incorrect. Ka for ammonium is based on the reaction
$NH_4^+ + H_2O$ <- -> $H_3O^+ + NH_3$
Now we have H^+.
We can calculate K_a from K_b.
$K_w = K_a * K_b$
$K_a = K_w/K_b = 1 \times 10^{-14} / 1.8 \times 10^{-5} = 5.555556 \times 10^{-10}$
So that
$Ka = 5.555556 \times 10^{-10} = [H_3O^+][NH_3] / [NH_4^+]$
and $5.555556 \times 10^{-10} = x^2 / (0.12-x)$
but we can approximate that 0.12-x is about equal to 0.12
(and then verify our assumption at the end)
$x = 8.16497 \times 10^{-6}$
pH = $-(\log [H^+]) = 5.09$

Diagnostic Feedback

Diagnostic Feedback for Response A:
Yes, indeed! Ka for ammonium is based on the reaction
$NH_4^+ + H_2O$ <- -> $H_3O^+ + NH_3$
Now we have H^+.
We can calculate K_a from K_b.

$K_w = K_a * K_b$
$K_a = K_w/K_b = 1 \times 10^{-14} / 1.8 \times 10^{-5} = 5.555556 \times 10^{-10}$
So that
$Ka = 5.555556 \times 10^{-10} = [H_3O^+][NH_3] / [NH_4^+]$
and $5.555556 \times 10^{-10} = x^2 / (0.12\text{-}x)$
but we can approximate that 0.12-x is about equal to 0.12
(and then verify our assumption at the end)
$x = 8.16497 \times 10^{-6}$
$pH = -(\log [H^+]) = 5.09$

Diagnostic Feedback for Response B:
Incorrect. Please check your work against the work below. Did you remember to change from K_b to K_a? We must do so on this problem in order to calculate pH. If we don't do that step, we end up calculating pNH_4^+.
Ka for ammonium is based on the reaction
$NH_4^+ + H_2O \leftarrow\rightarrow H_3O^+ + NH_3$
Now we have H^+.
We can calculate K_a from K_b.
$K_w = K_a * K_b$
$K_a = K_w/K_b = 1 \times 10^{-14} / 1.8 \times 10^{-5} = 5.555556 \times 10^{-10}$
So that
$Ka = 5.555556 \times 10^{-10} = [H_3O^+][NH_3] / [NH_4^+]$
and $5.555556 \times 10^{-10} = x^2 / (0.12\text{-}x)$
but we can approximate that 0.12-x is about equal to 0.12
(and then verify our assumption at the end)
$x = 8.16497 \times 10^{-6}$
$pH = -(\log [H^+]) = 5.09$

Diagnostic Feedback for Response C:
Incorrect. Only a salt made from a strong acid neutralizing a strong base can produce a neutral pH. Ammonium nitrate is not that kind of salt. It comes from NH_4OH and HNO_3, a weak base and a strong acid.
K_a for ammonium is based on the reaction
$NH_4^+ + H_2O \leftarrow\rightarrow H_3O^+ + NH_3$
Now we have H^+.
We can calculate K_a from K_b.
$K_w = K_a * K_b$
$K_a = K_w/K_b = 1 \times 10^{-14} / 1.8 \times 10^{-5} = 5.555556 \times 10^{-10}$
So that
$Ka = 5.555556 \times 10^{-10} = [H_3O^+][NH_3] / [NH_4^+]$
and $5.555556 \times 10^{-10} = x^2 / (0.12\text{-}x)$
but we can approximate that 0.12-x is about equal to 0.12
(and then verify our assumption at the end)
$x = 8.16497 \times 10^{-6}$
$pH = -(\log [H^+]) = 5.09$

Diagnostic Feedback for Response D:

Incorrect. This cannot be the answer. Recall that when we're dealing with salts, we examine the parentage to determine whether that salt will make a sample of water acidic or basic. The parentage of NH_4NO_3 is NH_4OH and HNO_3, a weak base and a strong acid. And we recall that such salts produce acidic solutions (pH less than 7).

Ka for ammonium is based on the reaction

$NH_4^+ + H_2O$ <- -> $H_3O^+ + NH_3$

Now we have H^+.

We can calculate K_a from K_b.

$K_w = K_a * K_b$

$K_a = K_w/K_b = 1 \times 10^{-14} / 1.8 \times 10^{-5} = 5.555556 \times 10^{-10}$

So that

$Ka = 5.555556 \times 10^{-10} = [H_3O^+][NH_3] / [NH_4^+]$

and $5.555556 \times 10^{-10} = x^2 / (0.12\text{-}x)$

but we can approximate that 0.12-x is about equal to 0.12

(and then verify our assumption at the end)

$x = 8.16497 \times 10^{-6}$

$pH = -(\log [H^+]) = 5.09$

Appendix A.4

Consent Form

IRB# 2003-08-0071

Informed Consent to Participate in Research

The University of Texas at Austin

You are being asked to participate in a doctoral research study on the effects of instructional methods on learning. This form provides you with information about the study. The Principal Investigator (the person in charge of this research) or her representative will also describe this study to you and answer all of your questions. Please read the information below and ask questions about anything you don't understand before deciding whether or not to take part. Your participation is entirely voluntary and you can refuse to participate without penalty or loss of benefits to which you are otherwise entitled.

Title of Research Study:

Investigation of Feedback on Student Performance

Principal Investigator(s) (include faculty sponsor), UT affiliation, and Telephone Number(s):

Marilla Svinicki (faculty sponsor), UT Dept of Educational Psychology, 232-1777
Joseph Lagowski (advisor), UT Dept of Chemistry and Biochemistry, 471-3288
Deborah Rush Walker (researcher), UT Dept of Chemistry and Biochemistry, 471-7732

Funding source:

This research project is not funded; it is part of a doctoral research project.

What is the purpose of this study?

The purpose of this study is to examine the effects of feedback on performance in freshman chemistry at The University of Texas at Austin. We expect 150-1,000 participants over the Fall and Spring semesters.

What will be done if you take part in this research study?

As part of the class, you will have the following: homework assignments, online quizzes, and exams. Thus, you will complete these assignments whether or not you participate in the research study. Procedures being tested in this study include standard questionnaires about instructional effects and demographics (like date of birth, gender, etc.).

What are the possible discomforts and risks?

There are no known risks.

What are the possible benefits to you or to others?

There may or may not be direct benefit to you. We do hope that the information learned from this study will benefit other students taking freshman chemistry in the future.

If you choose to take part in this study, will it cost you anything?

If you choose to take part in this study, it will not cost you anything.

Will you receive compensation for your participation in this study?

No monetary compensation will be provided for participants. Since you are enrolled in freshman chemistry at The University of Texas, you may receive credit for quizzes you complete, depending on your course syllabus. Such credit will be distributed regardless of whether or not you participate in this study.

If you do not want to take part in this study, what other options are available to you?

You can stop participating at any time. If you wish to stop your participation in this research study for any reason, you should contact: Deborah Rush Walker at xanthe3355@yahoo.com. You are free to withdraw your consent and stop participation in this research study at any time without penalty or loss of benefits for which you may be entitled. Throughout the study, the researchers will notify you of new information that may become available and that might affect your decision to remain in the study.

In addition, if you have questions about your rights as a research participant, please contact Clarke A. Burnham, Ph.D., Chair, The University of Texas at Austin Institutional Review Board for the Protection of Human Subjects, 512/232-4383.

How will your privacy and the confidentiality of your research records be protected?

Authorized persons from The University of Texas at Austin and the Institutional Review Board have the legal right to review your research records and will protect the confidentiality of those records to the extent permitted by law. Otherwise, your research records will not be released without your consent unless required by law or a court order.

If the results of this research are published or presented at scientific meetings, your identity will not be disclosed.

Will the researchers benefit from your participation in this study?

The researchers will not benefit monetarily from your participation.

Signatures:

As a representative of this study, I have explained the purpose, the procedures, the benefits, and the risks that are involved in this research study:

____

Signature and printed name of person obtaining consent **Date**

You have been informed about this study's purpose, procedures, possible benefits and risks, and you have received a copy of this Form. You have been given the opportunity to ask questions before you sign, and you have been told that you can ask other questions at any time. You voluntarily agree to participate in this study. By signing this form, you are not waiving any of your legal rights.

____

Printed Name of Subject **Date**

____

Signature of Subject **Date**

____

Signature of Principal Investigator **Date**

References

Angell, George W. (1949). The effect of immediate knowledge of quiz results on final examination scores in freshman chemistry. *Journal of Educational Research, 42,* 391-394.

Ames, C. (1992). Classrooms: goals, structures, and student motivation. *Journal of Educational Psychology, 84(3),* 261-271.

Ames, C. & Ames, R. (1984). Systems of student and teacher motivation: toward a qualitative definition. *Journal of Educational Psychology, 76(4),* 535-556.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: students' learning strategies and motivation processes. *Journal of Educational Psychology, 80,* 260-267.

Ammons, R. B. (1956). Effects of knowledge of performance: a survey and tentative theoretical formulation. *Journal of General Psychology, 54,* 279-299.

Annett, J. (1969). *Feedback and human behaviour.* Baltimore, MD: Penguin Books.

Baron, R.M. & Kenny, D.A. (1986). The Moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.

Bilodeau E.A. (1969). *Principles of skill acquisition.* New York City, NY: Academic Press.

Butler, R. & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology, 73,* 210-216.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences, 2nd edition.* Hillsdale, NJ: Lawrens Erlbaum Associates, Inc.

Cole, R.S. & Todd, J.B. (2003). Effects of web-based multimedia homework with immediate rich feedback on student learning in general chemistry. *Journal of Chemical Education, 80(11),* 1338-1343.

Duell, O.K., & Schommer-Aikins, M.S. (2001). Measures of people's beliefs about knowledge and learning. *Educational Psychology Review, 13(4),* 419-449.

Dweck, C.S. (1986). Motivational processes affecting learning. *American Psychologist, 41(10),* 1040-1048.

Earley, P. C., Northcraft, G. B., Lee, C., Lituchy, T.R. (1990). Impact of process and outcome feedback on the relation of goal setting to task performance. *Academy of Management Journal, 33(1)*, 87-105.

Fakhreddine, F.H. (1999). *Intensive Chemistry Seminar, Group Ability Composition, and Students' Achievement*. Dissertation Treatise, The University of Texas at Austin, TX.

Gabel, D. (1999). Improving teaching and learning through chemistry education research: a look to the future. *Journal of Chemical Education, 76(4),* 548-554.

Gagne', R.M. (1965). *The Conditions of Learning*. New York, NY: Holt, Rinehart and Winston, Inc.

Gagne', R.M. (1992). *Principles of instructional design, 4th edition.* Austin, TX: Harcourt Brace Jovanovich College Publishers.

Goodman, J.S., Wood, R.E., & Hendrickx, M. (2004). Feedback specificity, exploration, and learning. *Journal of Applied Psychology, 89,* 248-262.

Hall, R.W., Butler, L.G., McGuire, S.Y., McGlynn, S.P., Lyon, G.L., Reese, R.L., Limbach, P.A. (2001). Automated, web-based, second-chance homework. *Journal of Chemical Education, 78(12),* 1704-1708.

Harackiewicz, J.M., Barron, K.E., Tauer, J.M., Carter, S.M., & Elliot, A.J. (2000). Short-term and long-term consequences of achievement goals: predicting interest and performance over time. *Journal of Educational Psychology, 92(2),* 316-330.

Hashweh, M.Z. (1996). Effects of science teachers' epistemological beliefs in teaching. *Journal of Research in Science Teaching, 33(1),* 47-63.

Herron, J.D. (1996). *The chemistry classroom: formulas for successful teaching.* Washington, DC: American Chemical Society.

Hofer, Barbara K. (1997). *The Development of personal epistemology: dimensions, disciplinary differences, and instructional practices*. Dissertation treatise, The University of Michigan, MI.

Ilgen, D., Fisher, C, & Taylor, M. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology, 64(4),* 349-371.

Jacoby, J., Mazursky, D., Troutman, T., & Kuss, A. (1984). When feedback is ignored: disutility of outcome feedback. *Journal of Applied Psychology, 69,* 531-545.

Johnson, D.S., Perlow, R., & Pieper, K.F. (1993). Differences in task performance as a function of type of feedback: learning-oriented versus performance-oriented feedback. *Journal of Applied Psychology, 23,* 303-320.

Jonasssen, D.H., Wilson, B.G., Wang, S., & Grabinger, R. S. (1993). Constructivist uses of expert systems to support learning. *Journal of Computer-Based Instruction, 20(3),* 86-94.

Judd, C.H. (1905-6). Practice without knowledge of results. *Psychological Review Monographs Supplement, 7,* 185-198.

King, P.M., & Kitchener, K.S. (1994). *Developing reflective judgment: understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco, CA: Jossey-Bass Publishers.

Kulhavy, R.W. (1977). Feedback in written instruction. *Review of Educational Research, 47(1),* 211-232.

Kluger, A.N., & DeNisi, A. (1996). The Effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119(2),* 254-284.

Linnenbrink, E.A. & Pintrich, Paul R. (2002). Motivation as an enabler for academic success. *School Psychology Review, 31(3),* 313-327.

Locke, E.A., & Latham, G.P. (1990). *A theory of goal setting and task performance.* Englewood Cliffs, NJ: Prentice Hall.

Midgley, C.; Kaplan, A.; Middleton, M.; Maehr, M.; Urdan, T.; Anderman, L.H.; Anderman, E.; Roeser, R. (1998). The Development and validation of scales assessing students' achievement goal orientations. *Contemporary Educational Psychology, 23*, 113-131.

Milner-Bolotin, Marina. (2001). *The Effects of topic choice in project-based instruction on undergraduate physical science students' interest, ownership, and motivation.* Dissertation treatise, The University of Texas at Austin, TX.

Moreno, R. (2004). Decreasing cognitive load for novice students: effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32,* 99-113.

Nakhleh, M.B. (1992). Why some students don't learn chemistry. *Journal of Chemical Education, 69(3),* 191-196.

Narciss, S. (2004) The impact of informative tutoring feedback and self-efficacy on motivation and achievement in concept learning. *Experimental Psychology, 51(3),* 214-228.

Nicholls, John G. (1984). Achievement motivation: conceptions of ability, subjective experience, task choice, and performance. *Psychological Review, 93(3)*, 328-346.

Nicholls, John G., Cobb, P., Wood, T., Yackel E., & Patashnick, M. (1990). Assessing students' theories of success in mathematics: individual and classroom differences. *Journal for Research in Mathematics Education, 21(2),* 109-122.

Perry, W.G., Jr. (1968). Patterns of development in thought and values of students in a liberal arts college: a validation of a scheme. Cambridge, MA: Harvard University. (ERIC Document Reproduction Service No. ED 024 315).

Pintrich, Paul R. (2000). Multiple goals, multiple pathways: the role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92(3),* 544-555.

Pintrich, Paul R., Marx, R.W., & Boyle, R.A. (1993). Beyond cold conceptual change: the role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63(2),* 167-199.

Pintrich, Paul R. & Schunk, Dale H. (1996). *Motivation in education: theory, research, and applications.* Columbus, OH: Prentice Hall.

Posner, G.J., Strike, K.A., Hewson, P.W., & Gertzog, W.A. (1982). Accommodation of a scientific conception: toward a theory of conceptual change. *Science Education, 66(2),* 211-227.

Pressey, S. L. (1950). Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. *Journal of Psychology, 29*, 417-447.

Qian, G., & Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text. *Journal of Educational Psychology, 87(2)*, 28-292.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82(3)*, 498-504.

Schommer, M. (1993). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology, 85(3),* 406-411.

Schommer, M. (1998). The influence of age and education on epistemological beliefs. *British Journal of Educational Psychology, 68*, 551-562.

Schommer, M., Crouse, A., & Rhodes, N. (1992). Epistemological beliefs and mathematical text comprehension: believing it is simple does not make it so. *Journal of Educational Psychology, 84(4),* 435-443.

Taber, K. (2002). *Chemical misconceptions: prevention, diagnosis, and cure. Volume 1.* London, England: Royal Society of Chemistry.

Taconis, R., Ferguson-Hessler, M.G.M, & Broekkamp, H. (2001). Teaching science problem solving: an overview of experimental work. *Journal of Research in Science Teaching, 38(4),* 442-468.

Taylor, M.S., Fisher, C.D., & Ilgen, D.R. (1984). Individuals' reactions to performance feedback in organizations: a control theory perspective. *Personnel and Human Resources Management, 2,* 81-124.

Thorndike, Edward L. (1927). The law of effect. *American Journal of Psychology, 39,* 212-222.

Trowbridge, M. & Cason, H. (1932). An experimental study of Thorndike's theory of learning. *Journal of General Psychology, 7,* 245-260.

VandeWalle, D., Cron, W.L., Slocum, J.W., Jr. (2001). The role of goal orientation following performance feedback. *Journal of Applied Psychology, 86(4),* 629-640.

Vygotsky, L.S. (1978). *Mind in Society: The Development of Higher Psychological Processes.* Cole, M., John-Steiner, V., Scribner, S., & Souberman, E., Eds. Cambridge, MA: Harvard UP.

Weiner, B. (1990). History of motivational research in education. *Journal of Educational Psychology, 82(4),* 616-622.

Wilkinson, W.K., & Maxwell, S. (1991). The influence of college students' epistemological style on selected problem-solving processes. *Research in Higher Education, 32(3),* 333-350.

Yeany, R.H., Waugh, M.L., & Blalock, A.L. (1979). The effects of achievement diagnosis with feedback on the science achievement and attitude of university students. *Journal of Research in Science Teaching, 16(5),* 465-472.

Yeany, R.H., Dost, R.J., & Matthews, R.W. (1980). The effects of diagnostic-prescriptive instruction and locus of control on achievement and attitudes of university students. *Journal of Research in Science Teaching, 17(6),* 537-545.

Yeany, R.H., & Miller, P.A. (1983). Effects of diagnostic/remedial instruction on science learning: a meta analysis. *Journal of Research in Science Teaching, 20(1),* 19-26.

Zoller, Uri (1990). Students' misunderstandings and misconceptions in college freshman chemistry (general and organic). *Journal of Research in Science Teaching, 27(10),* 1053-1065.

VITA

Deborah Rush Walker began her journey in this world on November 5, 1971, in Pampa, Texas, the daughter of Jerry Lane Rush and Anita Louise Rush. After graduating valedictorian from Caprock High School in Amarillo, Texas, she entered the Physics department at West Texas A&M University. She graduated in 1994 with a dual-major bachelor's degree in chemistry and physics. During her undergraduate career, she assisted with research in cancer and alternative energy as well as teaching undergraduate physics labs. She entered the Chemistry Department at the Texas A&M Graduate School in 1994. In 1997 she married her true love, Marcus Eugene Walker, and followed her heart in research, as well, making the switch from bench research to chemical education. Combining a love of teaching and technology, she simultaneously served as adjunct faculty at Austin Community College while completing a Master's degree in Chemical Education from Texas A&M University in 1999. Eager to apply her newfound knowledge, Deborah worked for HCI Training in Austin, Texas, as both instructional designer and project manager. Curious about how to best formulate online feedback for corporate and academic sectors, she entered the Chemical Education group at The University of Texas at Austin in 2000. Her first son, Joshua Lee Walker, was born in 2002.

Permanent Address: 13118 Lubbock Lane, Austin, TX 78729

This dissertation was typed by the author.