THE EFFECT OF
AN ANALYTIC TRAIT, TASK-SPECIFIC RUBRIC
WITH OR WITHOUT EXPLANATION ON
SELF-EFFICACY AND ACADEMIC ACHIEVEMENT

A Thesis in
Instructional Systems
by
Steven J. McGriff

© 2007 Steven J. McGriff

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2007
The thesis of Steven J. McGriff was reviewed and approved* by the following:

Kyle L. Peck  
Professor of Instructional Systems  
Thesis Adviser  
Chair of Committee

Francis M. Dwyer  
Professor of Instructional Systems

Barbara Grabowski  
Professor of Instructional Systems

Peggy Van Meter  
Associate Professor of Educational Psychology

Susan M. Land  
Associate Professor of Instructional Systems  
Chair of Graduate Program

*Signatures are on file in the Graduate School.
ABSTRACT

The purpose of this quasi-experimental non-equivalent control group study was to investigate the effect of three strategies for implementation of an analytic trait, task-specific rubric—no rubric, rubric given without explanation, and rubric given with explanation—on teacher credential candidates’ self-efficacy for and achievement in writing a lesson plan.

Research Questions

1. Do the three strategies for analytic trait, task-specific rubric use lead to differences in the means for posttask self-efficacy for writing a lesson plan?

2. Do the three strategies for analytic trait, task-specific rubric use lead to differences in the quality of lesson plans written by subjects in the three treatment groups?

Methods

The study was implemented in the Fall of 2005 and Spring of 2006 at a large, west coast, public university with 114 volunteer participants recruited from introductory courses for 5th year, teacher credential candidates in which lesson plan writing was first introduced. Nine intact classes taught by seven different faculty instructors were assigned to one control and two treatment groups, based largely on requirements expressed by the instructors. In each of the courses, the course instructor using two unpublished instruments collected data: self-efficacy for writing a lesson plan (pretask and posttask) and a posttask survey of participant perception. The lesson plan assignments were collected and scored by two independent raters in a blind-review process using the rubric that had been provided to students.
Results

The findings indicate participants had an overall high level of self-efficacy for writing a lesson plan, even before the introduction of the rubric. Using a reliability-corrected ANCOVA model significant differences were found for posttask self-efficacy, favoring the control group. No significant differences were found on achievement for writing a lesson plan. Participants’ perceptions of rubric use, as reported on the posttask survey, indicated agreement with the perception that the rubric provided explanations for how to do the lesson plan, helped understanding of the requirements, helped establish goals, and helped with self-evaluation.

Limitations. The reading of this study should be done with an understanding of the study’s limitations with respect to recruitment of participants, assignment of treatments, adherence to procedures, equivalence of raters, instrumentation, characteristics of participants, and the nature of the task. This study was conducted in a naturalistic educational context in which the participants and their instructors interacted with one another and the study materials in a manner that was influenced significantly by instructors’ philosophies of education, views on the design and use of rubrics, and beliefs concerning the characteristics of a high-quality lesson plan. Generalization of these research findings should be done with caution.

Implications

Although this research study did not offer significant support for the claims that rubrics can promote self-efficacy or student achievement, more research is needed in controlled settings. The effective implementation of a rubric use strategy in this study was hindered by the limitations of the study and as a result, seems to have exposed a set of conditions under which rubric use may interfere with learner self-efficacy. As a
result, it is recommended that instructors ensure congruence of the course grading policy with the use of the rubric. If a rubric is used as part of an instructional strategy, a period of time should be given to learners for practicing how to use the rubric to increase performance on the task.

Recommendations for future research are discussed.
TABLE OF CONTENTS

LIST OF TABLES.......................................................................................................................x

LIST OF FIGURES.....................................................................................................................xii

ACKNOWLEDGEMENTS ..............................................................................................................xiii

CHAPTER 1 INTRODUCTION ..........................................................................................1

Statement of the Problem........................................................................................................2
Purpose of the Study..................................................................................................................3
Research Questions..................................................................................................................3
Main Research Questions........................................................................................................3
Supplemental Research Questions..........................................................................................4
Null Hypotheses......................................................................................................................5
Generalizability of Research.................................................................................................5
Definition of Terms..................................................................................................................6

CHAPTER 2 REVIEW OF THE LITERATURE.............................................................................8

Rubrics......................................................................................................................................8
Rubrics as an Assessment Method..........................................................................................9
Benefits of Rubrics for Teachers...........................................................................................10
Instructional Benefits of Rubrics...........................................................................................11
Instructional Strategy: Explanation of Rubric Use.................................................................13
Designing Rubrics..................................................................................................................16
Self-efficacy.............................................................................................................................19
Self-efficacy and Motivation.................................................................................................19
Self-efficacy Versus Self-concept.........................................................................................21
Measuring Self-efficacy............................................................................................................25
Self-efficacy Summary............................................................................................................26
The Relationship of Rubrics to Self-efficacy.........................................................................27
Conclusion..............................................................................................................................28

CHAPTER 3 METHODS AND PROCEDURES.....................................................................30

Pilot Study.................................................................................................................................30
Participants...............................................................................................................................30
Procedures, Materials, and Instrumentation..........................................................................30
Results.....................................................................................................................................31
Changes Made Based on Pilot Study.....................................................................................33
Major Study...............................................................................................................................35
Limitations of the Study...........................................................................................................35
Description of Participants.....................................................................................................38
Description of Courses and Settings....................................................................................42
Description of Course Instructors..........................................................................................46
APPENDIX I.............................................................................................................187
  Overview of Research Procedure.................................................................187

APPENDIX J.............................................................................................................188
  Alternative Research Study...........................................................................188
    Explanation of this Alternative Research Study......................................188
    Title ...........................................................................................................188
    Purpose of Study.......................................................................................188
    Research Questions .................................................................................189
    Literature Review .....................................................................................189
    Research Design .......................................................................................189
    Sampling ....................................................................................................191
    Variables ....................................................................................................192
    Methods of Data Collection ......................................................................192
    Ethics and Human Relations ....................................................................193
    Timeline .....................................................................................................194
LIST OF TABLES

Table 2.1  Blank Format for a Typical Analytical Trait Rubric with 3 Levels.................17
Table 2.2  Outcome Expectations ........................................................................................23
Table 3.1  Split-half Correlation of Self-efficacy Instrument.............................................33
Table 3.2  Instructors, Courses and Semester of Participation............................................49
Table 3.3  Rater Description ............................................................................................56
Table 3.4  Instructor G Lesson Plan Assignment................................................................64
Table 3.5  Participants by Semester, Course, Instructor, and Treatment (n=114)............66
Table 3.6  Hypotheses, General Research Questions, Data Collection Methods, Instruments, and Data Analysis Methods ....................................................................................79
Table 3.7  Research Questions Based on the Survey, Data Collection Methods, Instruments, and Data Analysis Methods .....................................................................................80
Table 4.1  Course Settings..................................................................................................84
Table 4.2  Participants by Course and Treatment (n=114)................................................88
Table 4.3  Academic Major by Treatment .......................................................................90
Table 4.4  Descriptive Data for Level of Confidence to Get an “A” grade on Lesson Plan Assignment .........................................................................................................................93
Table 4.5  Descriptive Data for Self-efficacy Before Treatment (PreSE)..........................94
Table 4.6  Inter-rater Reliability Correlations...................................................................95
Table 4.7  Adjusted Descriptive Data for Achievement on Lesson Plan 1 (LP1) ............95
Table 4.8  Years of Teaching Experience by Treatment .................................................97
Table 4.9  Prior Experience Creating Lesson Plans by Treatment..................................98
Table 4.10 Prior Experience Using Rubrics to Evaluate Own Work..............................100
Table 4.11 Experience Using Rubrics to Evaluate Students’ Work...............................101
Table 4.12 Descriptive Data for Level of Confidence to Get an “A” grade on Lesson Plan Assignment .........................................................................................................................103
Table 4.13 Adjusted Descriptive Data for Self-efficacy After Treatment (PostSE).........103
Table 4.14 ANCOVA for Self-efficacy After Treatment (PostSE) Covarying PreSE ....104
Table 4.15 Descriptive Data for Achievement on Lesson Plan 2 (LP2).........................105
Table 4.16 ANCOVA for Lesson Plan 2 (LP2) ..................................................................106
Table 4.17  Chi Square Analysis of Time Spent Creating Lesson Plan by Treatment...

Table 4.18  Chi Square Analysis of Overall Usefulness of Rubric by Treatment...

Table 4.19  Chi Square Analysis of Item 9a Rubric Provided Explanation for How to Do the Lesson Plan...

Table 4.20  Chi Square Analysis of Item 9b Rubric Did Not Help Understanding of Requirements...

Table 4.21  Chi Square Analysis of Item 9c Rubric Helped Establish Goals for Creating the Lesson Plan...

Table 4.22  Chi Square Analysis of Item 9d Rubric Did Not Provide Understandable Feedback...

Table 4.23  Chi Square Analysis of Item 9e Rubric Did Not Save Time in Completing the Assignment...

Table 4.24  Chi Square Analysis of Item 9f Rubric Helped by Providing Information for Self-evaluation...

Table 4.25  Relationship Between Dependent Variables...

Table 4.26  Relationship Between Participants’ Perceptions of Using the Rubric...

Table 4.27  Summary of Findings for Supplemental Research Questions 1 and 2...

Table 4.28  Summary of Findings for Supplemental Research Questions 3a-3f...

Table 5.1  Hypotheses Tested...

Table 5.2  Summary of Supplemental Research Questions 1 and 2 with Interpretations...

Table 5.3  Summary of Supplemental Research Questions 3a-3f with Interpretations...

Table J.1  Alternate Research Study List of Variables...
LIST OF FIGURES

Figure 4.1  Likert-type Scale for Survey Item 9...............................................................110
ACKNOWLEDGEMENTS

Thank you my Lord, God! It is because of your grace and mercy that I had the patience, persistence, and perseverance to accomplish this work. My inspiration is You:

Even youths grow tired and weary, and young men stumble and fall; but those who hope in the LORD will renew their strength. They will soar on wings like eagles; they will run and not grow weary, they will walk and not be faint.

Isaiah 40:30-31 (New International Version)

[Jesus] replied, "Because you have so little faith. I tell you the truth, if you have faith as small as a mustard seed, you can say to this mountain, 'Move from here to there' and it will move. Nothing will be impossible for you."

Matthew 17:20 (New International Version)

Even though the journey’s long
And I know the road is hard
Well, the One who’s gone before me
He will help me carry on
After all that I’ve been through
Now I realize the truth
That I must go through the valley
To stand upon the mountain of God.

(Mountain of God lyrics by Third Day)

I now realize that I had to also climb the mountain while God moved it!

It has been a great blessing to receive help from all the excellent people God placed in my life. A big, huge, enormous thanks to my wife, Gloria, whose love and commitment in supporting me through this lengthy adventure was unwavering, and my boys, Sebastian, Ethan, and Justin, who have only known their dad as a student, until now.

Thanks to: my dad, Asberry D. McGriff (1915–2004), who said that education could make all the difference in the world; the unfailing, special encouragement of Aunt Inez (McGriff) Fortier; my sisters Denyse and CarolAnn, with their families; and all my church family, who’ve long wondered when I would become a bona fide Ph.D....Dave
Smith, Tom Peterson, Gary Rodriguez, Kris Eng, Rich Carlson, Cheryl Sizemore, Arnold Argau, James Wright, and so many others through the seven year journey.

It takes a village to develop one PhD and mine was filled with the best and most caring collection of individuals, starting with an outstanding committee of scholars: Kyle Peck, my advisor and chair, who showed consistent dedication to helping, along with the faithful service of Barbara Grabowski, Frank Dwyer, and Peggy Van Meter. Thank you, Ed Yoder, for a piece of valuable insight given at the crucial eleventh hour.

Thanks to the village at Penn State with the many colleagues who have “walked” before me: Heeyoung Kim (a true friend!), Younghoon Kim, Ikseon Choi, Jaison Williams, Charles XiaoXue Wang, Xun Ge, Valerie Dudley, Gus Prestera, and John Nworie. No job is finished until the paperwork is done, so many thanks are given to special Penn State angels who looked over me and handled the administrivia with care: Nancy Crowder, Sherry Corman, Suzanne Harpster, and a special, heartfelt gratitude to Joan Schumacher, former director of Awards and Fellowships.

I sincerely appreciate my colleagues at San José State University: Robertta Barba, my friend, mentor, and department chair for unfailing faith and encouragement; Dean Susan Meyers for her unfailing trust, and my faculty collaborators, Mei-Yan Lu, Stephanie Steffy, Brenda Fikes, Alexander Sapiens, Kristeen Pemberton, Cindy Sammet, John Higaki, Amy Strage. Thanks to my mentors, who were willing to walk alongside me through many challenges: Kevin Clark, whom I also call friend, and George Cochern, Professor Emeritus, SJSU.

There is no better measure of my thankfulness to all of you than the enduring remembrance of your love, support, encouragement, and generosity inscribed on my heart. Thank You, God for each and every one of these generous people.
CHAPTER 1
INTRODUCTION

Rubrics are among the popular innovations in education (Andrade, 2001, 2005; Arter & McTighe, 2001; Goodrich, 1996; Popham, 1997; Stevens & Levi, 2005). A rubric is an evaluation tool widely used by teachers to assess complex learning outcomes (Goodrich, 1996; Popham, 1997) and is commonly described as a document that articulates the expectations for an assignment by listing the criteria, or what counts, and describing levels of quality from exemplary to developing (Andrade, 2000, 2005). Current books and articles on classroom assessment are filled with claims about the potential for using rubrics to improve student learning and achievement. Research suggests that rubric use can be related to improvements in student projects, but the literature offers little explanation for this relationship or their effectiveness in promoting academic achievement (Andrade, 2001, 2005).

Self-efficacy is believed to be an important variable in understanding students' behaviors in educational contexts (Schunk, 1989) and to significantly influence students' motivation to learn and skillful performance or achievement (Pintrich & Schunk, 2002; Schunk, 1989). Albert Bandura defines self-efficacy as, “People’s judgment of their capabilities to organize and execute courses of action required to attain designated types of performances” (Bandura, 1986, p. 391). A person's level of self-efficacy is relevant for a particular task and may vary under differing circumstances (Bandura, 1997; Pajares, 1996). For example, the extensive work of Schunk (1984, 1989, 1991, 1996, 2003) has found students' self-efficacy in academic settings varies within the process of acquiring knowledge and skills. Variations are the result of prior experiences, social supports, and such personal qualities as abilities and attitudes. Some factors that have
been found to enhance learners’ self-efficacy include: skills for self-evaluation, self-monitoring, and experiencing progress towards set goals (Schunk & Ertmer, 2000); setting goals (Schunk & Ertmer, 2000; Zimmerman, Bandura, & Martinez-Pons, 1992); possessing knowledge of what to do, perceiving a challenge worth accomplishing, and receiving performance feedback about progress (Pintrich & Schunk, 2002); and the teaching of effective learning strategies that will help them accomplish the set goals (Corno & Mandinach, 1983; Schunk, 1987b).

These influential factors for self-efficacy are characteristic of the benefits claimed for using rubrics. The type of rubric most suitable for promoting self-efficacy is an analytic trait, task-specific rubric because it presents distinct levels of quality for each criterion of a set of criteria determined to be the standard measure of performance for the task.

Given the claims that rubric use can positively affect students’ academic achievement, has characteristics that can promote self-efficacy for a task, and that self-efficacy enhances academic achievement, it follows that rubrics could serve as a mediating factor of self-efficacy to raise academic achievement.

**Statement of the Problem**

Research is needed to investigate the claims that rubrics improve student achievement in light of the empirical findings that self-efficacy has been shown to promote student achievement (Pintrich & Schunk, 2002). Although the literature outlines the benefits of rubrics for the instructor and learner, few studies have provided specific evidence of the effectiveness of any rubric use strategy to learning outcomes. Augmenting rubric use with an instructional strategy has the possibility of increasing the learner’s self-efficacy for the task and consequently, academic achievement. In
particular, it is worth exploring whether the potential of rubric use as an instructional strategy implemented prior to the student starting the task is significant (Whittaker, Salend, & Duhaney, 2001).

**Purpose of the Study**

The purpose of this research study was to investigate the effect of three strategies for implementation of an analytic trait, task-specific rubric—no rubric, rubric given without explanation, and rubric given with explanation—on teacher credential candidates’ self-efficacy for and achievement in writing a lesson plan.

**Research Questions**

The following are the research questions investigated in this study. The two main questions of the research study are presented first, followed by supplemental questions.

**Main Research Questions**

1. Do the three strategies for implementation of analytic trait, task-specific rubric use (no rubric, rubric given without explanation, and rubric given with explanation) lead to differences in the means for post-task self-efficacy for writing a lesson plan?

2. Do the three strategies for implementation of analytic trait, task-specific rubric use (no rubric, rubric given without explanation, and rubric given with explanation) lead to differences in the quality of lesson plans written by subjects in the three treatment groups?
Supplemental Research Questions

1. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in the amount of time spent creating the second lesson plan?

2. Do the two strategies for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of the usefulness of the rubric?

3a. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of how to write a lesson plan before starting the task?

3b. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ understanding of the requirements to earn the highest possible score on the lesson plan assignment?

3c. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of the helpfulness of the rubric for setting goals to create the lesson plan?

3d. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of understandable feedback derived from using the rubric?

3e. Does the strategy for implementation of analytic trait, task-specific rubric use lead to different amounts of time saved completing the lesson plan writing assignment?

3f. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of using the rubric for self-evaluation while creating the lesson plan?
Null Hypotheses

$H_01$ There will be no significant difference in the means for post-task self-efficacy for writing a lesson plan among three treatment groups, no rubric, rubric given without explanation, and rubric given with explanation.

$H_02$ There will be no significant difference in the means for achievement for writing a lesson plan among three treatment groups, no rubric, rubric given without explanation, and rubric given with explanation.

Generalizability of Research

Generalization of these research findings should be restrained because of the limitations of the study presented in Chapter 3. Generalizations should be carefully considered before applying them to populations of learners or learning environments outside those described in this study. The participants involved in this study were primarily teacher credential candidates recruited from education courses designed to prepare students for the teaching profession. It may be concluded that the general findings are appropriate to similar types of students. The instructional strategy of explaining the rubric to learners should be applicable to the population of teacher credential candidates, educators, and instructors in a variety of content areas and educational contexts. The concept of increasing learner self-efficacy for a specific task through use of an instructional tool should be applicable in all content areas of education and training where learners with low self-efficacy are required to complete a complex task.
Definition of Terms

*Analytic trait, task-specific rubric.* Current use of the word “rubric” is not consistent with historic definition. The word *rubric* is derived from the Latin word, *ruber*, for red. In medieval times a rubric was a set of instructions or a commentary attached to a law or liturgical service and typically written in red. For that reason, *rubric* came to mean something that authoritatively instructs people. Currently, in educative assessment a rubric is a set of scoring guidelines for evaluating students’ work and typically contains a scale of possible points, descriptors for each level of performance, and sometimes indicators as examples for the descriptor performance (Stevens & Levi, 2005; Wiggins, 1998). Rubrics can also be described as a table of specifications of criteria or learning goals that sets levels of standards for performance and guidelines for a specific assignment (Andrade, 2000; Arter & McTighe, 2001). Although there are many types of rubrics, two broad categories define the most popular rubrics—analytic trait and holistic. A holistic rubric gives a single score or rating for the whole product or performance based on an overall impression of the student’s work. An analytic trait rubric separates a product or performance into key traits or criterion so that they can be judged independently (Whittaker, Salend, & Duhaney, 2001). A rubric that focuses on assessing a particular task, rather than a general category of tasks, is called a task specific rubric. The rubric used in this study is an analytic trait, task-specific rubric designed to evaluate a lesson plan (see Appendix C).

*Self-efficacy.* A person’s judgment of their level of confidence to plan and carryout the independent actions required developing a written lesson plan.

*Task.* The task was the course assignment to write a lesson plan on a topic chosen by the participant.
**Academic achievement.** The performance of the participant for writing a lesson plan, as evaluated and scored by two raters in a blind review process using the rubric. The measure of academic achievement is the total score obtained from the sum of eleven subscores related to the individual criteria for writing a lesson plan. Further description of the lesson plan assignment is provided in chapter 3.

**Explanation.** The explanation used in this study is the justification for and statements of information about using the rubric that were presented to participants using a lecture and discussion method of teaching. The explanation is comprised of verbal information supported by the visual representation of the same information, as is typically found in slide presentations.
CHAPTER 2
REVIEW OF THE LITERATURE

This chapter examines the literature on rubrics, self-efficacy, an instructional strategy of explanation of rubric use, and the potential relationship among these constructs.

Rubrics

A rubric is an evaluation tool widely used by teachers to assess complex learning outcomes (Goodrich, 1996; Popham, 1997) and is commonly described as a document that articulates the expectations for an assignment by listing the criteria, or what counts, and describing levels of quality from exemplary to developing (Andrade, 2000, 2005). Rubrics can also be described as a table of specifications of criteria or learning goals that sets levels of standards for performance and guidelines for a specific assignment (Andrade, 2000; Arter & McTighe, 2001). Arter and McTighe (2001) further categorize rubrics as two primary types, analytical trait and holistic, although many variations exist. A holistic rubric gives a single score or rating for the whole product or performance based on an overall impression of the student’s work. An analytical trait rubric separates a product or performance into key traits or criterion so that they can be judged independently (Whittaker, Salend, & Duhaney, 2001). A separate score is given for each trait and a total score for overall quality. A rubric that focuses on assessing a particular task, such as writing a lesson plan, rather than a general category of tasks, such as writing, is called a task specific rubric. The rubric used in this study (Appendix C) is an analytic trait, task-specific rubric designed to evaluate a lesson plan. In general, rubrics can be used for grading a wide variety of assignments and tasks: oral presentations, book reports, group work, research papers, lab reports, and portfolios, to
name a few. The analytic trait, task-specific rubric is especially well suited for evaluating complex tasks with multiple criteria. The rubric is organized by the criteria and levels of achievement that make it easy for the learner to identify the specific aspects of the task that are deemed important. The criteria and the levels of achievement of an analytic trait, task-specific rubric make it the best type of rubric to evaluate the affect on self-efficacy, which is influenced by the learner’s knowledge of the requirements for a task.

**Rubrics as an Assessment Method**

Rubrics are most commonly used as a grading tool, but are not suitable for all types of assignments, such as those that are simple, objective, well-defined, or assess one dimension of learning. Rubrics are best suited for assessing complex, subjective, open-ended, constructed response, and multi-dimensional products or performances for which the solution path is not clear to the learner (Goodrich, 1996; Stevens & Levi, 2005) and that require integration of multiple skills and criteria (Arter & McTighe, 2001). Arter and McTighe (2001) offer some insight on the use of a rubric as a grading, assessment tool. Student performance can be evaluated using two basic assessment methods: selected response/short answer (i.e., multiple choice, matching, true/false) and constructed-response (i.e., performance assessment, essay tests, projects). Constructed-response calls for students to show what they understand and can do, typically by constructing a tangible product or performing a demonstration. Constructed-response assessment is open-ended, resulting in a range of responses, reveals alternative strategies, shows varying degrees of sophistication, and presents different levels of proficiency. Student responses to constructed-response assessments cannot be scored using a single answer key or machine, rather human judgment is
needed. For that reason, this format has been labeled “subjective.” To help make these “subjective” judgments as “objective” or fair as possible, a consistent set of high-quality performance criteria that emphasize the most important aspects of performance are required (Arter & McTighe, 2001).

**Benefits of Rubrics for Teachers**

Educators look for tools and motivational processes that will enhance learners’ acquisition of knowledge, skills, and attitudes, as well as improve achievement of academic objectives. In classroom use, rubrics are tools to help teachers efficiently and quickly assess student work and are used at many academic levels to expedite the grading process (Arter & McTighe, 2001; Goodrich, 1996; Popham, 1997; Stevens & Levi, 2005; Wiggins, 1998). It is commonly believed that rubrics can provide teachers greater consistency and efficiency in scoring and help improve instruction. It is often difficult for teachers to assess important dimensions of performance for open-ended tasks, problem solving, reports, and complex performances with certainty. Clearly defined performance criteria communicate the important elements of quality in a product or performance by reducing subjective judgments used when evaluating student work (Arter & McTighe, 2001). Andrade (2005) notes that another benefit for teachers is clearly defined criteria and scoring guides that help clarify instructional goals and serve as instructional objectives. Educators who have scored student work using national or state standardized assessments for writing note that the process of internalizing a set of criteria, learning a common vocabulary, and scoring hundreds of samples of student work teaches them a great deal about the elements of quality writing (Arter & McTighe, 2001). For many raters, this assessment experience indicates they developed greater confidence in their ability to identify the features of effective writing. Subsequently,
raters return to the classroom more confident about their ability to teach students how
to create effective writing (Arter & McTighe, 2001).

**Instructional Benefits of Rubrics**

The instructional benefits of using any type of rubric go beyond teacher
efficiency for grading, to clarifying and communicating the learning objectives, and
promotion of principles of equity and fairness. Wiggins (1998) advocates the central role
of properly designed rubrics is not assessment, rather the improvement of performance
through learner self-adjustment. Andrade (2000, 2005) further elaborates that rubrics,
when correctly used, are effective teaching and grading tools providing benefits to both
teacher and student. The rubric’s purposes are to give students informative feedback
while working on the project and to give detailed evaluations of their final products
(Andrade, 2000, 2005; Stevens & Levi, 2005), which would otherwise be given by the
teacher, repeatedly to multiple students for the same assignment. Andrade and others
propose that rubrics improve student motivation and raise achievement when used in
conjunction with an instructional strategy, given to students before the task is started,
and are used to support learning during the process (Andrade, 2000; Stevens & Levi,
2005)

As an instructional tool, rubrics make teacher’s expectations very clear (Andrade
& Du, 2005); provide meaningful, informative feedback to learners about their strengths
and areas in need of improvement (Andrade, 2000); communicate the goals to students
and help them plan an approach to an assignment (Andrade & Du, 2005); support the
development of critical thinking, skills, and understanding (Andrade, 2001); and help
students develop personal habits of self-assessment (Saddler & Andrade, 2004; Stevens
& Levi, 2005), such as checking their work, revising, and reflecting on feedback
In a focus group study of undergraduate students, Andrade and Du (2005) found that undergraduate students’ perception of improvements in the quality of their writing (i.e., student achievement) was attributed to understanding and using a rubric. Another study by Andrade (2001) examined the effects of rubrics on 8th graders writing skills and found that simply handing out and explaining a rubric seemed to help students improve their writing. The findings of that study also indicate rubrics provide meaningful feedback for students and promote self-assessment; encourage critical thinking; communicate important academic skills and levels of quality; and support learners’ self-efficacy. The relationship between rubrics and self-efficacy will be discussed later in this chapter.

Meaningful feedback on student performance has been found to promote higher levels of academic achievement for all grade levels and is most effective when it contains as much detailed, specific information as possible (Brinko, 1993). In a review of research on feedback, Mory (1996) presents numerous descriptions of feedback in which the typical form is verbal from the instructor and directed specifically toward the student about his or her performance. Extending the definition, Pintrich and Schunk (2002) indicate that printed materials that inform the student of their performance can be used as a form of self-feedback. The performance criteria included in a rubric offers the student a description of the expected quality of a product or performance. Students benefit from the use of clearly defined performance criteria that encourage a shared vocabulary and a clearer understanding of the important dimensions of quality performance. The rubric offers information about the highest level of achievement possible, the strengths of their work and areas for improvement. As a result, the student can read the higher-level description as self-assessment guidelines for improvement. Clearly, self-guided feedback derived from a rubric that can only describe what a
performance or product should look like is not as effective as verbal feedback on how the student performed. Rubrics can be helpful in contexts where the instructor is not readily available to offer focused feedback and some form of self-guided evaluation is needed. Huba and Freed (2000) discuss that feedback that enables students to conduct self-assessment and self-improvement is a form of intrinsic motivation. Zimmerman and Schunk (2001) found in multiple research studies that self-assessment activities can promote achievement and inspire learning behaviors typically associated with academic success, such as goal setting, self-monitoring, and revision. A well-designed rubric makes self-assessment easier, particularly when used as a “goal-setting guide,” and shared with students prior to starting a task to help them understand the nature of the levels of achievement goals and performance expectations for each level (Andrade & Du, 2005; Arter & McTighe, 2001).

Rubrics encourage critical thinking by showing expected levels of performance or achievement and inspiring students to develop the reflective practice of self-assessment and self-improvement. In a study of 8th graders writing a persuasive essay, Andrade (2001) proposed that thinking-centered rubrics helped students think more deeply. To promote critical thinking, a criterion was included in the rubric that prompted the student to anticipate the other side of the argument and prepare an explanation in defense against it. Students without the rubric did not consider the counterpoint arguments against their claim.

**Instructional Strategy: Explanation of Rubric Use**

Used in conjunction with sound instructional practice, analytic trait rubrics can contribute to students’ development of higher order thinking skills such as, analysis, evaluation, and creativity (Huba & Freed, 2000). Some of the rubric’s criteria break
down the components of critical thinking in an explicit manner, while the descriptions of the criteria explicitly lay out the requirements of how the student’s execution of the criteria can be critically analyzed. Critical analysis of the rubric’s criteria may be so basic to instructors creating the rubric that they are left implicit in the assignment and so may be overlooked by the students until the assignment is complete.

One good instructional practice that promotes critical thinking and effective use of the rubric is explanation and classroom discussion of the rubric prior to the students beginning the assignment. Andrade discovered rubrics are not completely self-explanatory, so students need help in understanding rubrics and their use. When students were given a rubric without explanation, “The more motivated students anguished over what to do with it, and the less motivated filed it in their notebooks and promptly forgot about it” (Andrade, 2005, p. 29). The rubric and standards for success should not be a mystery to students, rather when students know the criteria in advance of their performance, they are provided with clear goals for their work. In the same way, when students have access to performance criteria and scoring guides, they have the opportunity to self-assess and improve their work as it is developed (Wiggins, 1998). This method can enhance the quality of student learning and performance, not just evaluate it (Arter & McTighe, 2001).

In practice, instructors who use a rubric often present it to the student without adequate explanation or discussion about its purpose or effective use for completing the assignment (Arter & McTighe, 2001; Stevens & Levi, 2005). General educational practice has shown that students who receive explanation on how to do an assignment or use a new instructional strategy have higher achievement than students who do not receive explanation. It follows that students who receive explanation for using an instructional tool, like a rubric, as part of a broader instructional strategy, should be able to
effectively use and apply that aid to improve academic achievement. Andrade (2000) notes that while research studies present evidence for a link between rubrics and student learning, the mechanisms behind the achievement advantage provided by rubrics are not known.

The instructional strategy of discussing the rubric and communicating its purpose in advance makes the instructor’s implicit expectations explicit. Interviews with undergraduates in an educational psychology course regarding rubric use revealed they liked having the rubric in advance so that they knew “what’s expected,” and contrasted it with the “guessing game” they felt they had to play when instructors did not provide a rubric (Andrade & Du, 2005). In addition, instruction on rubric use models, in reverse, the criteria by which the work will be graded (Stevens & Levi, 2005) and supports the development of academic skills and understanding pertinent to achieving a level of quality for the assignment (Andrade, 2000).

Analytic trait, task-specific rubrics allow instructors to communicate the importance of some academic skills over others. The criteria used in the rubric is assigned a value, points, or percentage, according to the importance of those criteria on the final product. Including a scoring value to a rubric clearly shows which components of the task most strongly relate to the desired learning outcome. The rubric designed for this study includes weighted factors for criteria in developing a lesson plan, such as, stated objectives, use of technology, and instructional strategies. These criteria are worth more in the overall grade than other skills and explicitly communicate what is important in terms of producing a quality product (Stevens & Levi, 2005).

Explanation on the use of a rubric to complete a task is believed to increase the overall utility and effectiveness of the rubric (Andrade, 2000; Andrade & Du, 2005; Stevens & Levi, 2005). Further, when explanation is a part of an instructional strategy at
the time the students receive the assignment, it is more likely to serve its purposes of informing students of the standards of performance and supporting their achievement (Andrade, 2001). When learners better understand performance standards for the task, it is likely they will be able to compare their performance with the standard (Andrade & Du, 2005). Research has found that this type of evaluative process can potentially motivate learners (2001) and lead the learner to create a product that meets a higher standard described in the rubric than they would have otherwise created (Andrade & Du, 2005; Zimmerman & Schunk, 2001). It follows that if students know how a rubric is correctly and effectively used and apply their knowledge, then academic achievement is likely to improve.

**Designing Rubrics**

In the most basic format, a holistic rubric may be as simple as a checklist of requirements for the learner to meet. A simple analytic trait rubric includes a task description, criteria for the task, and descriptions of 3 or 4 levels of achievement for each criterion. The more useful format of an analytic trait, task-specific rubric is composed of four basic parts specifying the scope of the assignment: a task description (the assignment), an evaluation scale (levels of achievement or performance), the criteria or dimensions of the assignment (a breakdown of the skills/knowledge needed in the assignment), and descriptions of what comprises each level of performance (specific feedback). The format and layout of the rubric is a grid, similar to that shown in Table 2.1.
A rubric depends on standards for specifying how well criteria must be met for a given performance level. Standards are typically established on patterns of actual performances and instances of exemplary performance. Criteria for differing levels of performance vary on a continuum from fully to incompletely. Student achievement on these levels of performance can be judged against two types of standards: absolute and developmental (Wiggins, 1998).

Absolute standards are established automatically by the description of the top score on the rubric and the specific work examples selected to anchor the rubric to an excellent standard. Further qualification of absolute standards leads to developmental standards, which indicate that excellent performance is based on the standards for particular cohorts. For example, in athletics, the best performance models have different

Both absolute and developmental standards are necessary to inform students about real excellence, similar to the work of young athletes that is ultimately judged against standards, models, or examples set by professionals (Arter & McTighe, 2001; Wiggins, 1998).

The best rubrics follow logic and chronology. Logic guides the development of the overall criteria against which the rubric should be evaluated and modified. The chronology of events in the design and development process is replicated in the sequence of criteria represented in the rubric (Wiggins, 1998). Types of rubrics include either holistic or analytic trait. A holistic rubric has only one general descriptor for performance as a whole. An analytic trait rubric has multiple rubrics corresponding to each dimension of performance being scored. For example, in writing lesson plans, separate “mini-rubrics” may be created for “objectives”, “technology used”, and “instructional strategies” and for “showing all steps” and “understanding of physical laws” in physics. Rubrics may cover a range of specificity: generic (very general), genre, subject, or task specific (narrowly defined). Task-specific rubrics address highly specific criteria derived from a particular task.

Drawing from the strategies presented in the literature (Andrade, 2000; Stevens & Levi, 2005; Whittaker, Salend, & Duhaney, 2001; Wiggins, 1998) one can design a rubric using an analytic approach. Like other assessment instruments, rubric design should consider the appropriateness and validity of the performance task, the criteria, and the discrimination of the descriptors. Given a valid task, does the rubric focus on features that really indicate quality performance or product? Have the most important and revealing dimensions of performance been identified, given the most appropriate
criteria for the achievement target? Is it written clearly with definitions, indicators, and samples of work? Does the rubric provide an effective way of discriminating between levels of performance? Do teachers and students find it useful for instruction and assessment? Does the rubric lead raters to score consistently? Is it fair to all students? A rubric should permit valid inferences about performance to the degree that what is being measured (scored) is what is important to performance, not just easiest to see and score (Wiggins, 1998).

Reliability is the counterpart to validity. The rubric enables consistent scoring across judges and time. The best designed rubrics allow reliable scoring when the evaluative language (“exemplary,” “developing”) and comparative language (“better than,” “worse than”) are changed to highly descriptive language that guides judges to distinguish the salient and distinctive attributes of each performance level (Wiggins, 1998).

**Self-efficacy**

Self-efficacy is a widely researched belief system proposed by Albert Bandura, who defines it as, “People’s judgment of their capabilities to organize and execute courses of action required to attain designated types of performances” (Bandura, 1986, p. 391). Research that builds on Bandura’s original concept is extensive (e.g., Bandura, 1997; Pajares, 1996; Pintrich & Schunk, 2002; Schunk, 1987b, 1991, 2003; Zimmerman & Schunk, 2001).

**Self-efficacy and Motivation**

Self-efficacy is a significant factor in motivation and has been shown to affect choice of activities, effort, and persistence (Pintrich & Schunk, 2002). People having weak, or low self-efficacy for accomplishing a task may avoid it and by contrast, those
who believe they have the capabilities are likely to engage in the task. The difference between the two is especially acute when they encounter difficulties. Efficacious students work harder and persist longer than those who hold doubts. In addition, efficacious students are characterized as task-focused and strategic thinkers, who attribute failure to insufficient effort, which supports a success orientation (Bandura, 1997). They perceive difficult tasks as challenges to be mastered instead of threatening circumstances to be avoided. This type of affirmative orientation generates interest, deep involvement in activities, encourages setting challenging goals and maintaining strong commitment to them. People acquire information to determine self-efficacy from their actual performance (e.g., self-feedback or feedback from others), vicarious experiences (e.g., observing others), forms of persuasion, and physiological symptoms (e.g., heart rate, sweating, muscle tension) (Bandura, 1997).

Research on self-efficacy reveals a variety of correlations with motivation, outcome expectations, and achievement and as a result, efficacy beliefs are important, active contributors in facilitating performances (Pintrich & De Groot, 1990; Pintrich & Schrauben, 1992; Pintrich & Schunk, 2002; Schunk, 2003; Schunk & Ertmer, 2000; Zimmerman, Bandura, & Martinez-Pons, 1992; Zimmerman, 1989). The breadth of literature states people with similar skills demonstrate different performance on the same task because of variation in perceived self-efficacy for the task. Perceived self-efficacy is governed by what the individual believes they can do with their skills under a variety of circumstances, not the number of skills possessed. Therefore, different people with similar skills, or the same person in a different circumstance, may perform below standard, satisfactorily, or extraordinarily, depending on fluctuations in their beliefs of personal efficacy (Bandura, 1997). When self-efficacy perceptions are high, individuals will attempt tasks that foster skill development and improve their
capabilities. By contrast, when self-efficacy is low, people will not take on new tasks that might help them learn new skills (Bandura, 1997).

Schunk (1987b) reports on a study that found students perceived learning from television to be easier than learning from print. Students who held higher self-efficacy about learning from television invested less cognitive effort in learning. When learning from written materials, higher self-efficacy led to greater cognitive effort. Students were more likely to be cognitively engaged in learning when the task was perceived as difficult, but less likely to exert effort and less cognitively engaged when the task was deemed easy. In a related example, people who doubt their ability to perform in particular domains of activity avoid difficult tasks in those domains. They cannot effectively motivate themselves, and they slacken their efforts or give up quickly when they encounter obstacles. In addition, they have low aspirations and weak commitment to the goals they chose to pursue (Pintrich & Schunk, 2002).

**Self-efficacy Versus Self-concept**

Pintrich and Schunk (2002) discuss that self-efficacy should not be confused with the similar ideas of task-specific self-concept and self-perceptions of competence. The definition of self-efficacy states, “to be able to organize and execute courses of action,” which includes a specific and situational view of perceived competence by including the behavioral actions or cognitive skills that are required for competent performance in a given domain (Bong & Clark, 1999; Pajares, 1996; Pajares & Miller, 1995). Self efficacy is more than mere self-recognition of being good in school, but rather the specific judgments of having the skills for particular tasks in the academic domain, such as correctly subtracting three-digit numbers or writing different types of research papers.
Self-efficacy further separates itself from self-concept and self-competence in that it is used in reference to some type of goal, as stated in the definition, “attain designated types of performance.” This is due to the situational perspective of self-efficacy, which does not support generalization across other areas (Pintrich & Schunk, 2002). One implication of including a goal orientation is that self-efficacy judgments for similar tasks may vary as a function of personal or environmental differences. For example, professional baseball pitchers may lower their self-efficacy for pitching a good game because they did not get enough rest since their last outing. A student’s self-efficacy for learning a particular topic in biology may be lower because the difficulty of the material to be learned is greater than the material learned earlier in the course.

**Outcome expectations**

Outcome expectations are a second construct related to motivational behavior and affect, and are typically related to self-efficacy. An outcome expectation is also related to the expectancy for success construct from expectancy-value theories. An example derived from Bandura (1986) explains that a baseball player who believes he can hit a game-winning homerun is making an efficacy judgment; the anticipated team recognition, cheers, news interviews, and self-satisfactions for such a performance comprise the outcome expectations. In an academic setting, students would have efficacy judgments about their capabilities, skills, and knowledge to master school-related tasks, but also have outcome expectations about what grades they might receive on the tasks.

Although outcome expectations and self-efficacy do not have to be related, the former are frequently related to the latter. How an individual behaves largely determines the actual outcome and similarly, beliefs about outcome expectations are
dependent on self-efficacy judgments (Bandura, 1986). In academic settings, students’ self-efficacy beliefs are likely to be highly related to their outcome expectations, as shown in the cells in the top right and bottom left corners in Table 2.2.

<table>
<thead>
<tr>
<th>Table 2.2</th>
<th>Outcome Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self Efficacy</strong></td>
<td><strong>Low outcome expectation</strong></td>
</tr>
<tr>
<td><strong>High self-efficacy</strong></td>
<td>Social activism</td>
</tr>
<tr>
<td></td>
<td>Protest</td>
</tr>
<tr>
<td></td>
<td>Grievance</td>
</tr>
<tr>
<td></td>
<td>Milieu change</td>
</tr>
<tr>
<td><strong>Low self-efficacy</strong></td>
<td>Resignation</td>
</tr>
<tr>
<td></td>
<td>Apathy</td>
</tr>
<tr>
<td></td>
<td>Withdrawal</td>
</tr>
</tbody>
</table>

**Sources of self-efficacy**

Factors that can positively affect self-efficacy are characteristic of rubrics, such as setting goals for learning outcomes, setting standards for task performance, and providing feedback to the learner (Andrade, 2000; Andrade & Du, 2005). Interventions that have been found to enhance learners’ self-efficacy include: skills for self-evaluation, self-monitoring, and experiencing progress towards set goals (Schunk & Ertmer, 2000); ability to synthesize and analyze content (Meier, McCarthy, & Schmeck, 1984); setting goals (Schunk & Ertmer, 2000; Zimmerman, Bandura, & Martinez-Pons, 1992); receiving rewards for meeting set goals (Schunk, 1984); possessing knowledge of what to do, perceiving a challenge worth accomplishing, and receiving performance feedback about
progress (Pintrich & Schunk, 2002); and the teaching of effective learning strategies that will help them accomplish the set goals (Corno & Mandinach, 1983; Schunk, 1987b).

**Sources of academic self-efficacy**

The extensive work of Schunk (1984, 1989, 1991, 1996, 2003) on self-efficacy takes Bandura’s original research into the realm of academic skill learning. Schunk (1996) hypothesizes that as students begin the learning activity they differ in their self-efficacy for acquiring knowledge and skills as a result of prior experiences, social supports, and such personal qualities as abilities and attitudes. As they engage in activities their self-efficacy is affected by personal influences (goal setting, information processing) and by situational factors (rewards, feedback). These factors serve as cues from which they derive information on how well they are performing. Motivation and self-efficacy are enhanced when students perceive they are marking progress toward their goals and becoming more competent.

Self-efficacy is believed to be an important variable in understanding students’ behaviors in educational contexts (Schunk, 1989) and to significantly influence students’ motivation to learn and skillful performance (Pintrich & Schunk, 2002; Schunk, 1989). A learner’s self-efficacy beliefs are a key factor in motivation for accomplishing a specific task (Pintrich & Schunk, 2002), such as arithmetic division problems (Schunk, 1981) and writing performance (Meier, McCarthy, & Schmeck, 1984). Moreover, it is a generative characteristic and therefore can change depending on the task (Bandura, 1997). The development or cultivation of self-efficacy is effected by such varied factors as age, environment, or experiences, which continuously change. For example, Bandura (1997) describes school as an agency for cultivating self-efficacy by mediating children’s development of cognitive competencies and acquisition of knowledge. As children
master cognitive skills, they develop a growing perception of their intellectual efficacy. A strong sense of efficacy enriches a high level of motivation and academic accomplishments (Schunk, 1984). To operationalize this concept, consider that at the start of a task, students have different beliefs about their capabilities to acquire knowledge, perform skills, master the materials, and so on (Schunk, 1989). This initial sense of self-efficacy varies as a function of previous experience and personal characteristics and therefore can be promoted.

In addition to the amount of effort, the quality of effort as measured by the use of deeper processing strategies and general cognitive engagement of learning has been strongly related to self-efficacy (Pintrich & Schrauben, 1992). Junior high students with high self-efficacy were more likely to report using various cognitive and self-regulatory learning strategies (Pintrich & De Groot, 1990). Schunk conducted a series of experimental studies (1982, 1983, 1984, 1987a, 1996) that found students with high self-efficacy mastered various academic tasks better than did students with lower self-efficacy. These studies showed that efficacy was a significant predictor of learning and achievement, even after prior achievement and cognitive skills were taken into consideration.

**Measuring Self-efficacy**

Bandura (1997) cautions researchers attempting to predict self-efficacy that in order to increase the accuracy of the prediction of participants’ academic outcome to use specific criterial tasks that closely correspond to the participants’ performance. By contrast, global or generalized self-perceptions of competence yield ambiguous findings that do not clarify the contribution of self-efficacy beliefs to performance (Pajares, 1996). Prediction of academic outcomes is enhanced when self-efficacy and performance more
closely correspond or in other words, when the participants’ judgment of capability are reasonably precise and matched to a specific outcome (Bandura, 1986). The skills required to accomplish the performance described in the self-efficacy instrument should be clear to the participant (Pajares, 1996).

Other guidelines from Bandura (1997) adapted for this study include:

- Measure the relations between efficacy beliefs and action in close temporal proximity to more accurately reveal the causation.
- Self-efficacy assessment should measure participant’s beliefs to fulfill different levels of task demands within the domain selected for study.
- The scale for items should follow standard methodology, where individuals record the strength of their belief on a 100-point or 10-point scale, ranging in 10-unit intervals from a low value of no assurance, through intermediate degrees, to complete assurance.
- Items should be phrased in terms of *can*, a judgment of immediate capability, rather than *will*, a statement of future intention.

**Self-efficacy Summary**

Summarizing the literature on self-efficacy, research has shown it to be an important mediator of all types of achievement behavior. Self-efficacy is situation specific, leading researchers to measure self-efficacy in a context sensitive manner and for very particular tasks (e.g., two-digit subtraction problems with regrouping). As a result of the situational specificity, self-efficacy beliefs are assumed to be more variable, fluctuating, and changeable than the more static and stable self-concept and self-competence beliefs. Efficacy is not an inherent, fixed ability of knowing what to do and having the motivation to do it. By contrast, it is a generative capability in which
cognitive, emotional, and behavioral sub skills must be organized and coordinated to serve different purposes (Bandura, 1997). Individuals’ self-efficacy for a specific task on a given day might fluctuate due to their preparation, affective mood, physical condition (sickness, fatigue), as well as external conditions such as the nature of the task (length, difficulty).

The Relationship of Rubrics to Self-efficacy

It is possible that rubrics can promote students’ efficacy for a task by serving as the mediating factor for some of the self-efficacy enhancement interventions noted above. Although a review of the literature has revealed little empirical evidence that establishes a relationship between rubrics and self-efficacy, the literature does yield a body of research about three compatible pairs of concepts that can be used to investigate a potential relationship: rubrics and self-evaluation (or self-assessment); self-evaluation and academic achievement; and self-efficacy and academic achievement. Self-evaluation is the key construct that ties rubrics to self-efficacy. Rubrics can be used as a self-evaluation tool for learners (Stevens & Levi, 2005) and self-evaluation is one of the factors affecting self-efficacy (Zimmerman & Martinez-Pons, 1990). The learner can self-assess his or her performance, in part, on feedback from others (Bandura, 1997), themselves (Schunk, 2003), or materials (Pintrich & Schunk, 2002). Self-assessment has been shown to support academic achievement thus providing a bridge to link rubrics and self-efficacy to academic achievement and build upon the knowledge base.

Rubrics hold the potential to support and enhance self-efficacy. The process of self-evaluation of performance can be supported by the use of a rubric that specifically describes the criteria against which the learner compares his or her performance. Another link between self-efficacy and rubrics is when learners know they are making
progress and success towards a set goal based on information of performance levels detailed in the rubric. The learner can set goals for task performance based on the criteria explicitly stated in the rubric or derived from it. The criteria, as well as the organization of the rubric, can inform the learner of the processes and procedures for completing the task and help them determine “what counts.” Rubrics can be linked to self-efficacy via feedback the learner receives while completing the task (Pintrich & Schunk, 2002). Determining one’s self-efficacy requires information about one’s performance, which can be evaluated with the standards set in the rubric. Self-efficacy can be increased when learners are taught effective learning strategies. Instruction on using a rubric can be considered a learning strategy and if taught as guided practice, is likely to increase self-efficacy, and subsequently academic achievement.

**Conclusion**

Rubrics are a tool for teaching and authentic assessment of complex tasks or performances that require integrated skills in any education setting. When used as a part of an instructional strategy with explanation on rubric use before the student begins the task, a rubric should support the learning process and enhance academic achievement. Rubrics communicate objectives and criteria on task performance, provide meaningful feedback, promote self-assessment, and therefore, should support self-efficacy for the task. Rubrics encourage critical thinking and offer a way for learners to set goals for learning outcomes, read the standards for task performance (achievement), and offers feedback to the learner with respect to levels of quality. Requirements of self-efficacy include setting goals, understanding the standards for task performance, knowledge of what to do, and providing feedback to the learner. Feedback is important for developing efficacious behaviors because it provides knowledge of progress on the
task and offers the learner a means for self-evaluation. Explanation on the use of the
rubric informs the learner of the standards and requirements of the task and serves as
guided practice, both of which should increase self-efficacy. Based on standard
educational practice, effective explanation increases achievement, which has been found
to be an important mediator of self-efficacy. Rubric use with an instructional strategy,
like explanation, should enhance learner achievement for a variety of educational tasks
and performances.
CHAPTER 3
METHODS AND PROCEDURES

This chapter describes the design of the study, including the materials and instrumentation used. A pilot study was conducted in preparation for the major study. Changes were made prior to the major study based on the results from the pilot study. A description of the pilot study is presented prior to a description of the major study.

Pilot Study

The purpose of the pilot study was to conduct a formative evaluation of the instruments, materials, and procedures. The instruments were a self-efficacy measure and a rubric. The materials included a lesson plan assignment, a scripted PowerPoint presentation on using analytic trait, task-specific rubrics, and a post study survey.

Participants

Participants were 34 volunteers recruited from a convenience sample of two intact graduate courses in instructional technology at a large, west coast, public university. One class was assigned the control group (n=5) and the other to the treatment (n=29). The course met on different nights of the week. The majority of participants were in-service teachers (76%). Nearly one-third (32%) had 2-5 years of teaching experience and more than one-third (36%) had 6-10 years experience.

Procedures, Materials, and Instrumentation

At the start of the pilot study, participants in both groups received instructions about the assignment to write a lesson plan and were allowed to ask questions. The treatment group then received a 20-minute, scripted, PowerPoint presentation on the
use of the analytic trait, task-specific rubric for writing a lesson plan and a rubric to use when completing the task. The control group did not receive the presentation or a rubric. Before receiving the self-efficacy instrument, participants in both groups were told general information about the instrument as a measure of confidence to perform a task not related to lesson plan development. The 26-item self-efficacy instrument used a labeled semantic differential rating scale with interval points 10-100, with 10 meaning the lowest level of confidence “not sure,” to accomplish the task and 100 the highest, labeled “very sure.”

The self-efficacy instruments were collected from the participants before they left the class. Participants had one week to complete the task on their own, at which time the self-efficacy instrument was administered again and a post-study survey was completed. The survey contained reflective items about the participants’ responses to their particular task experience. The treatment group survey had additional items related to using the rubric.

Following the conclusion of the pilot study, copies of the participants’ lesson plans were distributed to two independent raters to conduct a blind-rating method using the rubric. The raters were trained in the use of the rubric for assessing the quality of the lesson plans. The raters’ scores were averaged for each lesson plan to compute an overall score.

Results

**Lesson plan achievement.** The mean of the control group (n=5) was 21.5 and the treatment (n=29) was 23.0 out of 33 total possible points. The blind review, inter-rater correlation was low (r = 0.2).
**Self-efficacy.** The mean self-efficacy scores were high for both groups (maximum score: 10). The control group pre-task self-efficacy ($M=9.57$, $SD=.55$) was nearly the same as the post-task measure ($M=9.38$, $SD=.84$). The treatment group also had similar scores for pre-task ($M=8.93$, $SD=.76$) and post-task ($M=9.10$, $SD=.82$). The correlations of self-efficacy scores within groups between pre-task and post-task measures were high for control ($r=.94$) and treatment ($r=.84$).

Because the results of the two main variables in the pilot study showed no significant differences, revisions were made in the design of the major study.

**Reliability of self-efficacy instrument**

Regarding the reliability of the self-efficacy instrument, the coefficient of internal consistency of the instrument was computed from the results of the pre-task and post-task administrations and the two coefficients compared. The control group ($n=5$) and treatment group ($n=29$) data were combined since the instrument was the same for both groups, yielding a total sample size ($n=34$). The 26 items of the instrument were subdivided into even and odd items for analysis using the split-half method. The scores of the two subsets were computed for each participant and then correlated. Correlations were run separately for both the pre-task and post-task measures to check consistency of results. The pre-task and post-task data were not combined into one meta set because the instrument was administered on two occasions. The reliability coefficient obtained represents only half the test, but reliability is related to the entire length of the test, so the Spearman-Brown prophecy formula was used to make a correction in order to obtain the reliability $r'(tt)$ of the entire test. As shown in Table 3.1, the coefficient of internal consistency is high ($r = .98$) for both the pre-task and post-task.
Table 3.1  Split-half Correlation of Self-efficacy Instrument

<table>
<thead>
<tr>
<th>Correlation of SE Pre-Task data</th>
<th>Spearman-Brown Prophecy formula</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EVEN (M)</td>
</tr>
<tr>
<td>split-half EVEN (M)</td>
<td>1</td>
</tr>
<tr>
<td>split-half ODD (M)</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>$r'(tt)_{pre} = 0.9796$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlation of SE Post-Task data</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVEN (M)</td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
<tr>
<td>split-half EVEN (M)</td>
</tr>
<tr>
<td>split-half ODD (M)</td>
</tr>
</tbody>
</table>

Changes Made Based on Pilot Study

Changes were made to the selected population, self-efficacy instrument, rubric, and procedure based on the results of the pilot study. The population in the pilot study was a convenience sample that comprised a vast majority of experienced teachers with more than 5 years teaching experience. It was assumed that the years of teaching experience would proportionately correspond to the level of experience writing or using lesson plans. The population was changed for the major study to include only teacher credential candidates with few or no years of teaching experience and who were enrolled in classes that first introduce lesson plan writing.

The original self-efficacy instrument had 26 items written in 13 pairs, one pair for each criterion of writing a lesson plan. The split-half correlation for reliability of both the pre-task ($r=.98$) and post-task measures ($r=.98$) were high. For the major study, one item from each pair was removed and three additional items were added to assess a
general level of confidence about writing a lesson plan (see Appendix B). The new item 14 asked about the participant’s confidence in meeting the instructor’s expectations for the lesson plan assignment, which is not particular to writing a lesson plan, but may be influential to the participants’ overall self-efficacy. New items were added to the beginning and end of the instrument, like bookends, to first assess participants’ general confidence in getting an “A” grade before reading the criteria for writing a lesson plan and then to estimate their score on a lesson plan assignment that used the criteria in listed in the instrument. More information about the development of the self-efficacy instrument is provided in the next section.

The analytic trait, task-specific rubric was modified for the major study based on the low inter-rater correlation ($r=0.2$) and feedback from the raters. Three criteria on the rubric were eliminated because of confusing language and redundancy with other criteria. The language of the descriptors were edited for consistency and to better reflect the terms and concepts with which teachers were more familiar (see Appendix C). In addition, weighting factors were added to explicitly show the user the relative value of one criterion versus another. A weighting factor was a multiplier of two or three times the level of quality, valued as 3 for exemplary, 2 for competent, and 1 for developing.

A significant change was made to the procedure of the pilot study where the researcher taught the lesson plan writing assignment. In the major study, it was not possible for the researcher to teach the lesson on using a rubric to help with lesson plan writing to every participant in the treatment group. A decision was made to have the course instructor teach the lesson following a script.
**Major Study**

The main purpose of this quasi-experimental non-equivalent control group study was to investigate the effect of three strategies for implementation of an analytic trait, task-specific rubric, no rubric, rubric given without explanation, and rubric given with explanation, on teacher credential candidates’ self-efficacy and achievement for writing a lesson plan. Specifically, the research questions were:

1. Do the three strategies for implementation of analytic trait, task-specific rubric use (no rubric, rubric given without explanation, and rubric given with explanation) lead to differences in the means for post-task self-efficacy for writing a lesson plan?

2. Do the three strategies for implementation of analytic trait, task-specific rubric use (no rubric, rubric given without explanation, and rubric given with explanation) lead to differences in the quality of lesson plans written by subjects in the three treatment groups?

**Limitations of the Study**

The reading of this study should be done with an understanding of the limitations with respect to recruitment of participants, assignment of treatments, procedures, raters, instrumentation, characteristics of participants, and nature of task. This study was conducted in a naturalistic educational context in which all the participants and their course instructors interacted with one another and the study materials in a manner that was as much a part of the established course curriculum as possible. This may have introduced extraneous variables outside the control of the study, since implementation was influenced significantly by instructors' philosophies of
In order to find a large enough pool of participants who would write lesson plans, a decision was made to recruit from nine intact courses in the College of Education. Random assignment of participants to treatments was not possible. In theory, a true randomization would have been to randomly assign individuals into the treatments. It was believed the use of intact courses would provide enough incentive for participation and limit the rate of attrition.

Permission to recruit participants was obtained from the course instructors, who volunteered to collaborate with the researcher under certain conditions. The conditions were that faculty retained control over instruction given to the participants and that some faculty agreed to participate in only certain treatment groups, based upon the amount of class time needed to conduct the study. The need for a high level of faculty collaboration was due mainly to the facts that the nine intact courses met at different times, different days, and in different geographic locations. Another limitation to implementing the study was that the lesson plan assignment occurred in each course at different times in the semester. Logistically, it was not possible for the researcher to conduct the study in every course on the date and time needed without assistance from the faculty members.

The assignment of intact courses was made based upon the condition of participation of the collaborating faculty member. Those faculty members who expressed commitment to presenting the 20-minute presentation on rubrics were assigned to the most time intensive treatment, rubric with explanation group. Faculty, who were willing to give the rubric to participants, but not the presentation, were assigned to the rubric without explanation group. Faculty who were willing to give
only the pre- and post-task self-efficacy instrument, were assigned to the control group. Treatment groups of at least forty participants were then organized based on the number of enrollments per class. This condition presented a challenge to internal validity, namely the effect of differential selection.

The procedures for implementing the study were carefully scripted and given to each faculty collaborator at each step of the study. However, control of instruction, nature and timing of feedback, and length of time to complete the lesson plan were retained by the faculty member.

The model of lesson plan presented in the rubric was created for this study based on the literature in instructional technology. It was not the model used by the course instructors, who most often used the Hunter 5-Step model. The reliability of the rubric was .83, however construct validity was not established.

In terms of scoring the lesson plans, 13 raters were employed in a blind-review process. Each rater was trained in the methods of a blind-review and instructed in techniques for maintaining consistency in scoring, but not all training occurred at the same time or under the same conditions.

Although the research procedure stated course instructors should use the rubric to score participants’ lesson plans, only one instructor reported she tried. It may be that course instructors did not use the rubric to score the lesson plans and that participants were aware that their grades in the course would not be affected by the rubric score.

The self-efficacy instrument developed for this study was based, in part on general designs used in research (Bong & Hocevar, 2002; Pajares, 1996) that had been shown reliable and valid. While reliability was determined to be .95, construct validity was not established.
The results should not be generalized to populations outside of teacher credential candidates, who may possess different educational and professional experiences and aptitudes.

Similarly, the results should not be generalized to different tasks that involve a set of cognitive skills and contextual knowledge outside the area of writing lesson plans.

**Description of Participants**

The study was implemented in the Fall of 2005 and Spring of 2006 at a large, west coast, public university with 114 volunteer participants recruited from introductory courses for teacher credential candidates in which lesson plan writing was first introduced. The university classified the participants as either fifth-year or post-baccalaureate students. Courses were in the departments of elementary education, child and adolescent development, and instructional technology. The elementary education courses included the participants’ first student teaching assignment. The child and adolescent development course was primarily composed of undergraduates interested in teaching and participants in the instructional technology courses were primarily teachers in their first year of teaching who were taking the course to complete the technology requirement to earn the clear credential. Participants were recruited in collaboration with the course faculty member. The rationale for using this particular participant pool and student profile was primarily to have participants with little or no experience in developing lesson plans, the central task in this study. In addition, the research topic and activity were relevant to the participants’ academic and professional interests. Potential benefits for participants included exposure to useful instructional
strategies and assessment methods that could be applied in their own professional practice in educational settings.

The University profile

Located in a major metropolitan region, this university is the West Coast’s oldest public university. Founded in 1857 as a teacher’s college, today, the university has a basic Carnegie Classification of “Masters L,” a large master’s degree institution and was comprised of 30,000 students, 2,400 faculty and staff members and 134 degree programs on a 154-acre downtown campus. In the Fall 2005 semester, when this study was conducted, the university enrollment was 54% women and 46% men comprising 74% undergraduates. 21% graduate, and 3.5% teacher credential candidates. The percent of students aged 19 years and under: 15%; students aged 20-24, 44%; 25-29, 20%; 30 and over, 21%. The ethnic profile was African American 4.5%, American Indian/Alaskan 0.4%, Asian 30.9%, Filipino 7.3%, Mexican American 10.7%, Other Hispanic 4.1%, Pacific Islander 0.9%, White 27.5%, and not stated, 13.8%. The geographic breakdown shows 53% of students enrolled are residents of the county, 40% from other counties in the state, and 6% from out of state, including international locations.

The College of Education profile

The College of Education was comprised of eight academic departments with a total of 197 faculty members who served 2,509 students in the Fall 2005, representing 8.3% of the university enrollment. The college enrollment comprised 23% undergraduate and 77% graduate students. The total number of teacher credentials granted by the college was 1,018.
**Academic major or program**

The participants were matriculated in different academic majors and programs described below.

Elementary Education, Multiple Subject Credential Program is the way to earn a license to teach in California schools, primarily at the elementary, middle or junior high levels. The Multiple Subject Teaching Credential authorizes its holder to teach in a "self-contained" classroom in which the same teacher teaches several subjects to the same group of students. Most elementary schools are organized around self-contained classrooms. The Multiple Subject Teaching Credential is also an authorization to teach in middle school "core programs" and, with relevant supplementary authorizations, in subject oriented departments (e.g., math or science).

Students in Multiple Subjects may combine their program with other Elementary Education Credential Programs. Students who applied for the program were required to have a grade point average (GPA) for the last 60 units of approximately 2.87; earned a passing score on California Basic Education Skills Test (CBEST); completed the U.S. Government and Constitution Requirement; completed a pre-professional experience in a public school setting or approved field-based course in the major; passed the California Subject Examinations for Teachers (CSET) Test; earned a bachelor's degree; received a Certificate of Clearance (This document verifies that the individual has completed the fingerprint and character and identification process and has been cleared by the Commission to begin student teaching.); and demonstrated basic skills in technology.

The following three specialized Teacher Credential programs offer further options for the student who has met the requirements noted above in Multiple Subjects.
Elementary Education, Critical Research Academy is a credential program that combines Multiple Subject Credential and Master of Arts program with an emphasis in Critical Research on Language, Culture, and Society. It offers a dynamic program for K-12 teachers primarily interested in education for social justice, development of cultural literacy, education to promote democracy, and pedagogy for teaching in urban schools. The program also incorporates requirements for both the Multiple Subject Teaching Credential and the Master of Arts in Education. This program starts in the Fall semester.

Elementary Education, School-University Partnership Internship (SUP) Program is a two-year internship in which successful candidates are compensated for full-time classroom teaching while satisfying the requirements for the Multiple Subject Teaching Credential with an optional Spanish BCLAD emphasis (available for interns who qualify). This option may be started only in the summer.

Elementary Education, The Spanish BCLAD (Bilingual Cross-Cultural, Language and Academic Development) Program is available to candidates wishing to teach in a bilingual classroom. Students in this program are required to demonstrate competence in Spanish and Latino Culture. The BCLAD option may be combined with other Credential Programs.

Some participants were enrolled in the College of Education, Department of Child and Adolescent Development (ChAD), Bachelor of Arts, Plan A (Subject Matter Preparation). ChAD leads to a baccalaureate degree (BA) and subject matter preparation for entry into a Multiple Subject Teaching Credential Program after graduation. Plan A is relevant to a wide range of careers in education. Under this plan, coursework is diversified to cover five content areas typically taught in K-8 classrooms: language and literature; mathematics and science; humanities and the arts; human development and physical education; and history and the social sciences.
Some participants were enrolled in the College of Humanities and Arts, Liberal Studies, Bachelor of Arts with emphasis on teacher preparation. Objectives of the program prepare the student to: engage in critical inquiry and exploration: interpretation, analysis, synthesis, argument and expression (written and verbal); demonstrate knowledge of available information resources; describe significant ideas, structures and values in the disciplines that underlie the K-8 curriculum: history, social science, mathematics, science, reading, language, literature, visual and performing arts, physical education, health, and human development; systematically examine connections among the major subject areas commonly taught in K-8 (conceptual foundations, content, and underlying values); investigate key ideas related to the California Student Academic content Standards and Curriculum Frameworks for grades K-8; and display understanding of the key issues and practices of K-8 schooling.

**Description of Courses and Settings**

The participants in this study were recruited from nine intact courses in the College of Education taught by seven faculty members during the academic year 2005-06.

**Elementary Education 143a, Orientation to Student Teaching (EDEL 143a):** The university course catalog states this semester long, 5-unit class covers the role of state and local government in education and offers clinical observation of classroom, school and district organization. The purpose was to help participant become familiar with duties/responsibilities of classroom teacher and begin to plan for and teach small groups and whole class with emphasis on lesson planning. The grading scheme was credit/no-credit. The class was scheduled for three days per week, four hours per class
Participants spent two days per week at their designated school site and one day in seminar classroom setting.

Five different sections of EDEL 143a were used in this study. The course section settings varied with each instructor based on either the school district in which the students were doing their student teaching or a particular area of academic specialization. The EDEL 143a course sections described below are identified by an instructor code used in place of the instructors’ names (i.e., Instructor A, B, C, D, and E). Descriptive information for all instructors is provided in more detail after this section.

EDEL 143a taught by Instructor A met off-campus on Friday mornings in a high school classroom, approximately 12 miles from the university. The total enrollment for his Fall 2005 course was 7 students, which allowed for the course to be taught as a seminar with the instructor facilitating discussion and information dissemination. The students most likely selected this section because of its proximity to their assigned student teaching sites.

Instructor B met her EDEL 143a students off-campus on Tuesday mornings at various locations that included restaurants, elementary school meeting rooms, and the County Office of Education meeting. These locations were within 12 miles of the university campus. The total enrollment for Spring 2006 was 12 students and the course was taught in seminar style allowing for students to discuss course topics with the instructor as facilitator. The students most likely selected this section because of its proximity to their assigned student teaching sites.

Instructor C taught her EDEL 143a course 4 miles off-campus on Friday mornings in a large-group, meeting room located at an elementary school district office. The total enrollment for Fall 2005 was 24 students who sat clustered around conference tables in groups of eight for primary instruction and discussion facilitated by the
instructor. The students most likely selected this section because of two reasons; its proximity to their assigned student teaching sites and its designation for students enrolled in the Critical Research Academy.

Instructor D taught his EDEL 143a course off-campus Thursday evenings at an elementary school library located approximately 8 miles from the university. The total enrollment for Fall 2005 was 18 students. The instructor used a round-table, seminar setting to facilitate discussion and lead learning activities. The students most likely selected this section because it offered the BCLAD emphasis.

Instructor E taught her EDEL 143 course off-campus Wednesday mornings at an elementary school district office located approximately 4 miles from the university. The total enrollment for Fall 2005 was 27, taught in seminar style with students clustered in table groupings around the room. The instructor lead discussion and presentations from the front center of the room. The students most likely selected this section for its specialization in handling the School University Partnership (SUP) program option.

The four remaining courses used in this study were from other College of Education departments, Child and Adolescent Development and Instructional Technology.

**Child and Adolescent Development 149, Child Health and Physical Activity (ChAD 149):** This semester long, 3-unit course is described in the university course catalog as an integrative approach to understanding the multiple factors that impact children’s health and physical fitness and that influence the development of lifelong habits. Emphasis was on the inter-relationships between health and fitness and cognitive, social, and emotional well-being. ChAD was a major designed for students interested in teaching in elementary school or middle school. This Fall 2005 class met on campus in a large, lecture hall with stadium seating and a stage for two days per week
(Mondays and Wednesdays), 65 minutes per session and had a total enrollment of 33 students. Instructor F led the learning experience from the stage primarily in lecture format.

**Instructional Technology 122, Microcomputers in Education (EDIT 122):** This semester long, 3-unit course is described in the university course catalog as using microcomputers in an educational environment and application programs (word processing, databases and spreadsheets), as well as other software for classroom management and instruction. The purpose is for preparing, teaching and evaluating a lesson integrating existing software. EDIT 122 is a first level, technology integration course for teachers seeking the Professional Clear Credential, which must be completed within 5 years after obtaining the preliminary teaching credential. The course met on Monday afternoons in an on-campus, computer lab setting and included a 3 hour combined lecture presentation and hands-on lab. Total enrollment for this section of EDIT 122 was 12. Instructor G used lecture, small group interaction, and individual skill development strategies.

**Instructional Technology 241, Emerging Technologies for Teachers (EDIT 241):** This semester long, 3-unit course is described in the university course catalog as a systematic review of results of research studies in techniques of designing, selecting, producing, using and evaluating the use of curriculum materials and instructional media in teaching; research studies in mass communication media; and procedures applicable to setting up small-scale evaluative studies of curriculum materials and media programs within schools. EDIT 241 is a second level technology integration course for teachers seeking the Professional Clear Credential. The combined class and lab met on-campus once per week for 3 hours per session. Two sections of EDIT 241 were used in this study: the Fall 2005 section with total enrollment of 20 students met
on Tuesday afternoons and the Spring 2006 section with total enrollment of 15 students met Wednesday evenings. Instructor G, who used lecture, small group interaction, and individual skill development strategies, taught both sections.

**Description of Course Instructors**

The instructors for the courses used in this study were also collaborators in the study. Each instructor allowed access to their students for recruiting and handled the exchange of study materials and instruments between the researcher and the participants. The following description of the instructors provides background information of their position in the university, educational experience, and additional background information.

**Instructor A**: Lecturer, Elementary Education, instructor for EDEL 143a. He had over 5 years college teaching experience for this course at the time of this study. His background includes 21 years as a high school social studies teacher and counselor, 1 year as a high school associate principal, 8 years as a district level director of personnel, followed by 5 years as the principal of an adult education, Welfare to Work program before assuming the position of lecturer at the university. Instructor A earned a doctorate in curriculum and instruction from Brigham Young University.

**Instructor B**: Lecturer, Elementary Education, instructor for EDEL 143a. Instructor B had 35 years of education service in K-6 school settings, including 17 years as an elementary principal. She reported her greatest personal accomplishment was being the principal of a school that received the National Blue Ribbon from Bill Clinton in Washington, D.C. Prior to going into administration, Instructor B was an E.S.L. Resource Teacher and worked closely with Dr. Elizabeth Cohen at Stanford University.
on research about science as a means to teach E.S.L. to students. Currently retired, Instructor B has been a lecturer for two years at the university.

**Instructor C:** Lecturer, Elementary Education, instructor for EDEL 143a. Educational background includes BA, Elementary Education; BA, Fine Arts; MA, Learning Disabilities; and Ed.D, Educational Leadership. Teaching experiences include 8 years as an elementary school teacher and 12 years as a resource specialist, and 5 years in adult education and workplace settings. Instructor C has taught in the EDEL program for 15 years, the last six as a full time lecturer, covering credential courses (Classroom Management, Action Research, Social Studies, Multicultural Foundations, and Student Teaching Supervision) and master’s level courses (Sociology of Education, Critical Studies in Language, Culture, and Narrative) that include serving as MA Project and Thesis Advisor. She is currently involved in a federally funded project investigating case studies of college teaching in light of the principles of universal design.

**Instructor D:** Professor, Elementary Education, instructor for EDEL 143a. Education background includes a M.A. and Ph.D. in Education from Stanford University. Instructor D has been on faculty for twenty years and during his tenure served as Field Placement Director for student teachers.

**Instructor E:** Lecturer, Elementary Education, instructor for EDEL 143a and current Field Placement Director for student teachers. Instructor E has a doctorate in education and has been on faculty for 13 years, preparing candidates for the teaching profession.

**Instructor F:** Professor, Child and Adolescent Development, instructor for ChAD 149. For the past decade, she has been involved in research examining the effects of infusing service learning across the curriculum. Instructor F has received a federal teacher quality enhancement grant to recruit and prepare math, science and special
education teachers. Instructor F also earned national recognition for co-directing the “It Takes a Valley” project, a federally-funded program designed to prepare future teachers to serve in high-need schools. Since joining the faculty in 1987, Instructor F has been involved in recruiting and supporting college students who are interested in careers in education. Her efforts in this area have allowed her to work closely with school, community, business and industry partners to create a wide variety of hands-on learning experiences for future educators. She works to acquaint these teachers with the challenges they will face and the resources they can draw upon as they pursue their professional goals. Instructor F also conducts research in the areas of student motivation and service-learning. Her educational background includes a Ph.D. in Developmental Psychology from the University of California at Berkeley.

Instructor G: Professor, Instructional Technology, instructor for EDIT 122 and EDIT 241. Instructor G joined the faculty in 1988 after earning her Ph.D. from Arizona State University. In the late 1990s, she was a Senior Fellow in the School of Communication Studies at Nangyang Technological University in Singapore. Instructor G has been teaching in the public schools, universities and consulting for government agencies, educational institutions, business and industry in the United States and overseas for the last 20 years. She is an internationally recognized author and instructional designer and has been invited to give keynote speeches, seminars, workshops in major conferences, universities, and companies in the United States, China, Taiwan, Hong Kong, Macau, Singapore, Thailand, and Germany. Her area of specialization includes: interactive multimedia design/development/evaluation, teaching strategies, using multimedia to enhance thinking and creativity, designing multimedia for international audience, teacher training/support, and using educational technologies to help English Language Learners.
Table 3.2 summarizes the information about faculty collaborators and their courses with respect to the particular semester in which they participated in the study.

### Table 3.2 Instructors, Courses and Semester of Participation

<table>
<thead>
<tr>
<th>Semester</th>
<th>Instr.</th>
<th>Course</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2005</td>
<td>A</td>
<td>EDEL 143a (Orientation to Student Teaching)</td>
<td>7</td>
</tr>
<tr>
<td>Spring 2006</td>
<td>B</td>
<td>EDEL 143a (Orientation to Student Teaching)</td>
<td>12</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>C</td>
<td>EDEL 143a (Orientation to Student Teaching)</td>
<td>24</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>D</td>
<td>EDEL 143a (Orientation to Student Teaching)</td>
<td>18</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>E</td>
<td>EDEL 143a (Orientation to Student Teaching)</td>
<td>27</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>F</td>
<td>ChAD 149 (Child Health and Physical Activity)</td>
<td>33</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>G</td>
<td>EDIT 122 (Microcomputers in Education)</td>
<td>12</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>G</td>
<td>EDIT 241 (Emerging Technologies for Teachers)</td>
<td>20</td>
</tr>
<tr>
<td>Spring 2006</td>
<td>G</td>
<td>EDIT 241 (Emerging Technologies for Teachers)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total Enrollments</td>
<td>168</td>
</tr>
</tbody>
</table>

**Description of Instruments**

**Self-efficacy Instrument**

The self-efficacy instrument used in this study was an unpublished instrument consisting of 16 Likert-type items designed to measure the perceived confidence level of the participant to plan and implement some of the criteria required for developing a lesson plan (see Appendix B). The instrument was developed for this study because an established self-efficacy instrument for assessing lesson plan writing was not found. The instrument was intended to differentiate high and low levels of self-efficacy among participants.
The basic structure of the self-efficacy for writing a lesson plan instrument was developed from guidelines written by Bandura (2005), based on good models of instruments identified by Pajares (1996, 2006), and in particular, the criteria derived from standards for lesson plans written in textbooks on instructional technology (Gagné, Wager, Golas, & Keller, 2005; Morrison, Ross, & Kemp, 2004; Smith & Ragan, 1999).

A self-efficacy instrument created by Varney (2003), titled the *Dissertation Appraisal Inventory*, provides the model for how the instrument created for this study was formed. Varney broke down the complex task of writing a dissertation into 16 subtasks and used a self-referenced, confidence response scale with three labels at equal intervals from 0 to 100, “no confidence at all,” “moderate amount of confidence,” and “completely confident.” In a similar manner, the task of writing a lesson plan was broken down into 14 subtasks, or criteria listed in the instrument used for this study. The response scale was based on one developed by Bong and Hocevar (2002), where participants recorded the strength of their belief on a 10-point scale, ranging in single-unit intervals from a low value of no assurance, through intermediate degrees, to complete assurance, for example, 1 corresponded to “not sure”; 4, “somewhat sure”; 7, “pretty sure”; and 10 corresponded to “very sure”. The guidelines for constructing self-efficacy response scales confirm that single unit intervals are an acceptable format so long as the number of intervals retain sensitivity and people are able to differentiate between intervals (Bandura, 2006).

The two-page instrument began with a set of instructions to establish the appropriate mindset about rating their strength of belief in their confidence to develop a lesson plan. A single response item was included on page 1 to gauge the participant’s level of confidence that they would obtain an ‘A’ grade on the lesson plan assignment.
The purpose of this initial item was to gain a quick response of participants’ overall level of confidence for writing a lesson plan. On the second page, the first 14 items focused on the different criteria that comprised a lesson plan. Following these fourteen items, the last item asked participants to circle the estimated score they would receive on a 100-point scale with 10-point increments. The scale for this item was changed to correspond to a standard grading scale and capture the participants’ confidence for writing a lesson plan after reading the criteria for a lesson plan and evaluating their level of confidence for successfully performing each one. Items 14 and 15 were crafted to gain a general reflection of participants’ confidence for meeting the instructors’ expectations and accomplishing the whole task.

Similar to the concept presented by Varney (2006), it was assumed that self-efficacy for writing a lesson plan could be calculated as the sum of the individual confidence levels a participant would have for completing all the subtasks or criteria associated with writing a lesson plan. The criteria for writing a lesson plan were compiled from textbooks in the field of instructional technology, as explained in more detail later below.

The total self-efficacy score was the average of participants’ ratings for items 1-14. The first and last items were not included in the total score because of the general nature of the item. Scores in the range 1-3 were categorized as low self-efficacy, scores of 4-6 as medium self-efficacy, and scores of 7-10 as high self-efficacy.

Reliability and validity

Internal consistency reliability estimates using Cronbach’s Alpha were .95 for both the pre-task and posttask administrations of the self-efficacy instrument. Validity of the instrument was checked using a face-value process with two professors of
education who reviewed the instrument and provided a verbal appraisal regarding its strengths and weaknesses for representing the factors of lesson plan design. Construct validity was not established. Reliability was measured in the pilot study and changes were made based on the results.

**Analytic Trait, Task-Specific Rubric**

An analytic trait, task-specific rubric is a scoring instrument that describes the performance criteria and levels of quality for the sub-components of a specific task, such as writing a lesson plan. In this type of rubric, each criterion (i.e., trait or subtask) of the lesson plan receives a separate score and then a total score is calculated by the sum of the sub scores, see Appendix C. By contrast, a holistic rubric is designed to provide one score for the entire product or performance.

As described in Chapter 2, the process for the analytic trait, task-specific rubric design involved analysis using questions posed by Wiggins (1998). Given a valid task, does the rubric focus on features that really indicate quality performance or product? Have the most important and revealing dimensions of performance been identified, given the most appropriate criteria for the achievement target? Is it written clearly with definitions, indicators, and samples of work? Does the rubric provide an effective way of discriminating between levels of performance? Do teachers and students find it useful for instruction and assessment? Does the rubric lead raters to score consistently? Is it fair to all students? A rubric should permit valid inferences about performance to the degree that what is being measured (scored) is what is important to performance, not just easiest to see and score (Wiggins, 1998).

To design a rubric with content validity, an analytic design process was used based on guidelines found in books and articles on rubrics (Stevens & Levi, 2005; Arter
First, the content of the rubric (the criteria or traits for a lesson plan) was derived from the literature in the field of instructional technology (Gagné, Wager, Golas, & Keller, 2005; Morrison, Ross, & Kemp, 2004; Smith & Ragan, 1999). Smith and Ragan (1999) suggest the lesson plan should be considered a document that describes what the instructor should do or say within the context of the learning goals and list the materials and resources necessary to successfully deliver the instruction. Gagné and colleagues (2005) suggest a lesson plan include four elements: a statement of the objective of the lesson; a list of the instructional events to be used; a list of the media, materials, and activities by which each event would be accomplished; and prescriptions for instruction that describe the roles and activities for each event. The lesson plan may be organized with the objective(s) at the top and a column for each of the other three items, noted above. Morrison and colleagues (2004) emphasize the message design of a lesson plan should be clear, well organized, and easy for someone else to follow.

The criteria for a lesson plan developed for the first version of the rubric used in the pilot study contained:

- Introduction
- Purpose
- Objectives
- Media equipment, materials, resources
- Technology used
- Lesson procedure
- Instructional events
- Instructional strategies (learning outcome, teacher activity, student activity)
- Evaluation
- Motivation
- Organization and presentation
Second, examples of lesson plans written by novice and experienced teachers were gathered and sorted into three groups (exemplary, competent, developing) based on how closely they matched the criteria from the instructional technology texts. Next, the important dimensions of performance (criteria) were analyzed among lesson plans in each grouping. General notes on the traits were organized using the same three groupings. The fourth step involved writing a description of each trait for the three levels of performance; exemplary, competent, and developing.

Reliability and validity

Internal consistency reliability estimates using Cronbach’s alpha was .83. Face validity of the rubric used in the pilot study was appraised by an instructional technologist (who was also an experienced elementary school teacher), who used an analytical “meta-rubric” described by Arter & McTighe (2001), to analyze four traits—content, clarity, practicality, and technical soundness. Each trait had a 3-point scale, where 3 indicated the rubric was “ready to roll,” 2 meant it was “on its way but needs revision,” and 1 equated with “not ready for prime time.” Each point on the scale was comprised of 4-8 descriptions of quality to guide the evaluator. A total score indicated overall quality of the rubric. The rubric used in the pilot study scored 3, “ready to roll.” Construct validity was not evaluated due to limitations of time.

Formative evaluation of the rubric was conducted after the pilot study by two independent raters. One rater had 15 years of elementary teaching experience and a recent master’s degree in instructional technology. The second rater had two years of elementary teaching experience. The raters suggested changes in the wording of the trait descriptions to clarify the distinction between the levels of performance. Two criteria, introduction and lesson procedure, were removed because it was felt they were
Redundant with other criteria. Weighting factors were added to emphasize the importance of some criteria (objectives, technology used, instructional strategies, evaluation, and motivation) over the rest. The scoring procedure was one commonly used for rubrics. Each criterion score was determined by the numerical value assigned for each performance level times the weighting factor. For example, exemplary performance had numerical value of 3, competent 2, and developing 1 and that value was multiplied by the criterion’s weighting factor, either 1 (no weighting factor), 2 or 3. The total score was calculated as the sum of the nine criterion scores for a maximum point score equal 51. The new scoring procedure was added to explain how the calculation with weighting factors was computed.

**Raters**

For the main study, thirteen raters each evaluated one or two sets of lesson plans using the rubric. Three raters participated in both the Fall 2005 and Spring 2006 implementations of the study. Each rater held a credential to teach elementary or secondary school. Eleven had earned a master’s degree in instructional technology. Each rater was paid a $100 stipend for their service. A brief description of each rater is provided in Table 3.3, followed by a description of the training procedure and then the rating procedure.
Table 3.3 Rater Description

<table>
<thead>
<tr>
<th>Rater Code</th>
<th>Description</th>
<th>Years Exp.</th>
<th>Semester</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM</td>
<td>Former 5th grade teacher, current home school teacher</td>
<td>10</td>
<td>Fall 2005, Spr 2006</td>
</tr>
<tr>
<td>MH</td>
<td>MA Instructional Technology, elementary school teacher</td>
<td>15</td>
<td>Fall 2005</td>
</tr>
<tr>
<td>ML</td>
<td>1st grade teacher</td>
<td>1</td>
<td>Fall 2005, Spr 2006</td>
</tr>
<tr>
<td>CO</td>
<td>MA Instructional Technology, elementary school teacher, English Second Language</td>
<td>4</td>
<td>Spr 2006</td>
</tr>
<tr>
<td>KF</td>
<td>MA Instructional Technology, middle school teacher, agricultural education</td>
<td>7</td>
<td>Fall 2005</td>
</tr>
<tr>
<td>VP</td>
<td>MA Instructional Technology, high school teacher, world cultures</td>
<td>4</td>
<td>Spr 2006</td>
</tr>
<tr>
<td>KM</td>
<td>MA Instructional Technology, 2nd grade teacher</td>
<td>2</td>
<td>Spr 2006</td>
</tr>
<tr>
<td>UF</td>
<td>MA Instructional Technology, 3rd grade teacher</td>
<td>5</td>
<td>Fall 2005, Spr 2006</td>
</tr>
<tr>
<td>CS</td>
<td>MA Instructional Technology, 1st grade teacher</td>
<td>12</td>
<td>Spr 2006</td>
</tr>
<tr>
<td>SC</td>
<td>MA Instructional Technology, elementary school teacher, technology specialist</td>
<td>17</td>
<td>Fall 2005</td>
</tr>
<tr>
<td>KH</td>
<td>MA Instructional Technology, middle school, language arts and social studies</td>
<td>6</td>
<td>Fall 2005</td>
</tr>
<tr>
<td>MS1</td>
<td>MA Instructional Technology, middle school, science, technology, engineering</td>
<td>12</td>
<td>Fall 2005</td>
</tr>
<tr>
<td>MS2</td>
<td>MA Instructional Technology, elementary school teacher</td>
<td>7</td>
<td>Fall 2005</td>
</tr>
</tbody>
</table>

**Rater training**

Raters were trained to review and score the participants’ lesson plans using a seminar format for 2-3 raters or one-to-one instruction. The seminar format allowed for better discussion and focused instruction that maximized trainer-rater interaction. The training design included nine instructional objectives using presentation, discussion, and practice, if needed. The researcher conducted the training sessions following a training guide he developed for the purpose (see Appendix H). Three sessions were conducted with groups of 2 or 3 raters and five sessions were conducted as one-to-one
training. The multiple sessions were given in order to accommodate the raters’ schedules.

The overall goal of instruction was that raters would learn the purpose of and procedure for evaluating lesson plans written by participants in the research study. In addition, raters would use the lesson plan rubric to effectively evaluate the level of participants’ achievement for writing a lesson plan.

The instructional objectives are included below.

**Objective 1:** Given instruction, the rater will be able to accurately describe the purpose of the research project and lesson plan evaluation process.

**Objective 2:** Given instruction, the participant will be able to correctly explain the purpose of inter-rater reliability and the role of the rater.

**Objective 3:** Given the rubric, the participant will be able to correctly identify the parts of the rubric (11 performance criteria, 3 performance levels, score points).

**Objective 4:** Given the rubric, the participant will be able to express the degrees of difference between the 3 performance levels in their own terms.

**Objective 5:** Given the rubric and a real lesson plan, participants will be able to correctly record a score for each performance criteria and calculate a total score for the lesson plan.

**Objective 6:** Given a set of scored sample lesson plans, the participant will be able to competently compare their scored lesson plan with the samples.

**Objective 7:** Given instruction, the participant will be able to competently execute the procedure for managing multiple lesson plan evaluations.

**Objective 8:** Given practice, the participant will be able to consistently record their notes on the evaluation process for each lesson plan.
Lesson plan assignments to raters

Faculty collaborators delivered the lesson plans created by participants before any markups or feedback were written on them. Each lesson plan was coded with a unique, three-digit number to identify it by faculty collaborator, course, treatment, and semester. Each course was assigned a number range, such as 101-149 for the first lesson plan and 151-199 for the second lesson plan. A participant’s first lesson plan was coded with the next available number in the allotted sequence. The participant’s second lesson plan was then coded with the number from the second range that was exactly 50 greater than the first. This method helped to quickly associate each participant’s two lesson plans. Raters did not have knowledge of the meaning of the coding scheme.

The sequence in which each course’s lesson plans were coded was based on the order in which the faculty collaborator submitted them. The participants name and course information were marked out to preserve confidentiality and minimize bias. 114 participants submitted two lesson plans and each was copied twice for a total of 456 lesson plans. The lesson plan copies were kept separate. The two sets of copies were both organized into matching stacks by the course and then further separated by first and second plans. Next, a rater’s set of 30 lesson plans was made by a haphazard process of selecting both copies of a participant’s lesson and making two identical sets. This process continued until 7 sets of paired lesson plans containing 30 lessons each and 1 set of paired lessons, each containing 18 were completed for a total of 16 sets of lessons. Haphazard means a lesson pair was sometimes selected from the top, middle, or bottom of different course stacks, but not in sequence. A rater’s stack typically contained an unequal mix of first and second lessons from various courses.

The 16 sets of lessons were distributed to 13 different raters, 3 of whom participated twice, once in each semester of the study. A set of lessons was given to a
rater during the training session. The corresponding matched set of lessons was
assigned to another rater who attended the next training session. Raters in the same
training session did not receive the same set as another rater in that session to reduce
the possibility of cross-rater collaboration during review. The number of lessons in a
stack (30) was determined by an estimate of how many lessons a rater could evaluate in
ten hours, the maximum amount of time raters stated they had available to work on this
study.

Survey

A post-study data collection survey was designed to gather participants’
demographic data, their response to the lesson plan task, and perceptions of using the
rubric (see Appendix A). The one-page, 15-item survey contained multiple choice
categorical data, Likert-type scales, and open-ended response items. In particular, items
2, 3, 5, 6, and 7, related to participants’ prior experiences. Items 4, 8, and 9a-f pertained
to conditions that were affected by the treatment. According to Thomas (1999), to
prevent patterned responses, the response choices for items 5 and 6 were reversed so
the first choice options, “always” and “frequently” respectively, were the highest level
of frequency response options for the item. For the same reason, items 9b, 9d, 9e, were
written in the negative tone. Instructions specified that all participants should respond
to items 1-7, but only those who received the rubric were also instructed to answer
items 8 and 9a-f.

The survey was used to gather a richer set of data that would hopefully provide
some insight to the results of the treatment conditions. Given that the task of writing a
lesson plan occurred outside the observation of the researcher, the survey was the
selected method to collect data about participants’ activity and reaction to the study.
Validity was determined by a face value process in which two professors of instructional technology reviewed each survey item to determine whether it would yield a response appropriate for the content being measured; the format to determine readability and organization of the items on the page; and the sequence of items to reduce the bias one response may impose upon a later response. Verbal feedback was recorded in notes written by the researcher. The process was completed twice and the survey was modified after each round.

Description of Materials

Rubric Presentation and Script

A 20-minute, scripted PowerPoint presentation on how to use the analytic trait, task-specific rubric to complete the lesson plan task was developed (see Appendix D). It included explanation on how the rubric could be used as a self-assessment tool and the potential benefit for helping the participant produce a higher quality (i.e., higher scoring) lesson plan. The presentation was designed for the faculty collaborator in the rubric with explanation group to use for information dissemination in an expedient manner using the direct instruction method. The format and timing of the presentation was designed for the most efficient use of the instructor’s limited time, the capacity of the adult learner to learn from a verbal-visual, lecture-format presentation, and fit the technological advantages afforded in the classroom.

The presentation was designed as a Microsoft PowerPoint slideshow with 17 slides and printed speaker notes that explained the definition and benefits of a rubric, the parts of a rubric (e.g., performance criteria, performance levels, descriptions,
scoring,) as well as explanation of how to use the rubric for the lesson plan assignment. The speaker notes were printed with the slide image for the instructor to read to the participants. In addition, a set of instructions for how to use the presentation and speaker notes were reviewed with each instructor prior to implementation.

**Lesson Plan Assignment**

Each faculty collaborator taught lesson plan design differently. All the Elementary Education courses (EDEL 143a) introduced participants to the Madeline Hunter 5-step model (anticipatory set, presentation, guided practice, independent practice, and closure), but each faculty collaborator set different expectations for which model the participant was to use. For example, four representative lesson plan assignments from Instructors A, B, C, and G, are shown below. The lesson plan assignments from Instructors D, E, and F were not made available for inclusion here.

Instructor A reported he used the Hunter 5-step approach for planning his own lessons, but that the candidates were free to choose another model for their lesson plans. He explained his rationale was that the lesson plan was not the critical factor in evaluating a candidate’s teaching effectiveness. In fact, he did not evaluate the lesson plan; rather he observed the lesson implementation as presented in the candidate’s classroom context. Given this was the candidate’s first semester of student teaching, his evaluation of their performance included verbal and written feedback on the following seven items:

1) Did the teacher credential candidate speak loud enough so the student in the back could hear?

2) Did the candidate notice whether a student was off task?
3) If they did notice, how did the candidate help the distracted student get back on task?

4) Did the candidate teach the objective as listed in the lesson plan?

5) Was the candidate prepared? For example, did they have enough materials? Did they check to make sure the technology worked? Did they know what they are teaching?

6) Did the candidate demonstrate good classroom management skills?

7) Was the candidate aware of the time?

In a similar way, Instructor B modeled the Hunter 5-Step approach, taught the approach, and expected her candidates to use it in their own lesson plans. Handouts on the Hunter 5-Step approach were given to candidates along with additional information from Robert Marzano that highlighted the Hunter 5-Step approach. She also discussed brain theory as it related to the Hunter 5 Step approach. Candidates sent Instructor B an electronic copy of their lesson plan 48 hours prior to the scheduled classroom observation of the lesson implementation. She reviewed the plan with the candidate before observing the lesson and offered feedback to the candidate with respect to the Hunter 5 Step approach. The lessons submitted for this study were unedited and did not contain any of Instructor B’s comments or included any revisions made later by the candidate.

Instructor B stated that a good lesson plan follows the Hunter 5 Steps with a very clear and specific objective related to the state standards. During observation, she would look for transitions that are under one minute. In addition, it was expected that higher order thinking questions based on Bloom’s taxonomy would be incorporated into the plan and asked at the appropriate time in the lesson.
Instructor C taught lesson planning in her EDEL 143a course in a more elaborate manner. First, she introduced the backward design model of lesson planning to emphasize the importance of focusing on how credential candidates would know whether their elementary students had learned the lesson objectives. Five models were presented in the course (inquiry, sheltered instruction, cooperative learning, 6-step, and skills development) with activities designed to help candidates recognize the appropriate use of each model. For writing the lesson plan, Instructor C required candidates to use the 6-step method because the candidates’ cooperating teachers also knew that model. Cooperating teachers were required to collaborate with candidates in developing the lesson plan. Candidates brought their completed plans to class for peer review and debriefing with Instructor C. The second lesson plan was created to fulfill the requirement for a reading/language arts teacher, and became part of the candidate’s coaching cycle. Feedback on this lesson plan was provided in various ways, including individual conference, written comments, scored against a checklist, and by using the rubric provided in this study. The Instructor C lesson assignment is included in Appendix E.

Instructor C described a good lesson plan design for beginning student teachers as one that included plenty of details concerning:

- key questions that guide instruction
- timing
- clear links between objectives and assessment
- differentiating strategies that clarify the candidates’ expectations of marginal students (including gifted, hyperactive, English Language Learners and learning handicapped)
- lists of materials and resources needed
The lesson plan was the candidate’s mental rehearsal and a reflection of their knowledge of their students and the content they were teaching. In addition to this required information on the lesson plan, Instructor C required candidates to include a paragraph on how the lesson connected with the planned curriculum for that grade level.

By contrast, Instructor G in the instructional technology course (EDIT 241) used a simplified approach to lesson plan instruction. A lesson plan format should include content standards, preplanning, the lesson described step-by-step, assessment, and extensions to support the lesson. The assignment is shown in Table 3.4.

<table>
<thead>
<tr>
<th>Table 3.4 Instructor G Lesson Plan Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lesson Plan</strong></td>
</tr>
<tr>
<td>Use the Lesson Plan Format (or similar) given in class to develop a plan for implementing a virtual field trip.</td>
</tr>
<tr>
<td><strong>Supporting Materials</strong></td>
</tr>
<tr>
<td>1) Virtual Field Trip Guide. Include a set of directions to help students navigate the site. The directions should include colored photographs, graphics, and text (taken from the tour site or other related source) to increase student interest and keep them on the prescribed path.</td>
</tr>
<tr>
<td>2) Student activity. Design an activity the students will complete during or after the virtual field trip (e.g., scavenger hunt, writing assignment, crossword puzzle, project, etc.) to reinforce the major concepts learned as a result of the trip.</td>
</tr>
</tbody>
</table>

Description of Treatment Groups

Rubric with explanation (RE)

This group received the rubric and a 20-minute, scripted slide presentation that presented information in lecture and discussion format on a description of the rubric.
and how the rubric could be used as an instructional tool for feedback and self-assessment while writing a lesson plan. The assignment for writing the lesson plan was created by the course instructor and given as part of the normal curriculum. The content of the lesson plan was of the participant’s own choosing or based on a content domain selected by the course instructor. The lesson plan was created in a paper-based format.

**Rubric without instruction (R)**

This group received the rubric, but without any explanation of its use. Like the rubric with explanation group, the course instructor as part of the normal curriculum gave the lesson plan assignment.

**Control (C)**

The control group received a lesson plan assignment by the course instructor as part of the normal curriculum without any rubric or explanation on the use of a rubric to complete the lesson plan task.

**Procedures**

**Matching Participants to Groups**

Intact classes of participants were matched to groups based on the instructor’s degree of commitment to the study. Participants were invited to volunteer and sign informed consent forms after hearing a brief description of the study. Accounting for both administrations of the study, the number of participants who signed informed consent forms were 50 for the control group, 54 for rubric without explanation, and 51 for rubric with explanation, for a total of 155 possible participants. Because of attrition over the duration of the study in each semester, 41 participants were excluded to yield
the final sample population of 114. The distribution of participants among treatment
groups is shown in Table 3.5.

Table 3.5  Participants by Semester, Course, Instructor, and Treatment (n=114)

<table>
<thead>
<tr>
<th>Semester</th>
<th>Instr.</th>
<th>Course</th>
<th>Control (C)</th>
<th>Rubric without explanation (R)</th>
<th>Rubric with explanation (RE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2005</td>
<td>A</td>
<td>EDEL 143a</td>
<td>4</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10.26</td>
<td>50.0</td>
<td>38.46</td>
</tr>
<tr>
<td>Spring 2006</td>
<td>B</td>
<td>EDEL 143a</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>C</td>
<td>EDEL 143a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>D</td>
<td>EDEL 143a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>E</td>
<td>EDEL 143a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>F</td>
<td>ChAD 149</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>69.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>G</td>
<td>EDIT 122</td>
<td>4</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11.0</td>
<td>39.0</td>
<td>17.95</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>G</td>
<td>EDIT 241 (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring 2006</td>
<td>G</td>
<td>EDIT 241 (2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>39</td>
<td>36</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Overview of Procedure

The following overview of the procedures provides a brief outline of the
sequence of the study. Detailed procedures for each experimental group follow in the
next section. The start of the research procedure for each treatment group varied in the
semester depending on the course schedule, which was set by the course instructor. The
end of the procedure was typically 2-3 weeks before the last class date of the semester.
See Appendix I for an overview schematic of the key phases of the procedure.
General procedure steps

1. **Initial contact**
   a. **Informed Consent:** Volunteer participants signed consent forms before first lesson plan assignment (Lesson Plan 1) was given.

2. **Lesson Plan 1:** At a later date, participants started, completed, and received feedback from the instructor on the first lesson plan assignment.

3. **Treatment for Lesson Plan 2, Start:** On the day the second lesson plan (Lesson Plan 2) assignment was introduced, participants received:
   a. **Self-efficacy (pre-task):** Participants completed the self-efficacy instrument
   b. **Treatment** based on their experimental group

4. **Complete Lesson Plan 2:** Participants worked independently for a period of time specified by their instructor to complete the lesson plan, typically 2-3 weeks.

5. **Lesson Plan 2, Finish:** On the day the second lesson plan was due,
   a. Participants submitted the lesson plan assignment
   b. **Self-efficacy (post-task):** Participants completed the self-efficacy instrument
   c. **Survey:** Participants completed the post-study survey

**Detailed Procedure for Each Treatment Group**

The following procedure for the study was designed to establish the initial contact for all groups:

1. Presented opportunity to volunteer in the study in the participants’ course.
2. Participants in the study were informed that the research study was an analysis of lesson plans that they would create in the semester as part of their normal curriculum, not something extra.

3. Participants were informed that the lesson plans they created would be reviewed by the researcher and that they would be asked to complete additional instruments.

4. Participant signed one informed consent form and returned it to the envelope. Participant was instructed to keep one copy of the form for their records.

5. Informed consent procedure completed.

6. At a later date, the first lesson plan assignment was completed and collected for data analysis. Lesson Plan 1 was scored and evaluated by the instructor and returned to the participant with feedback before Lesson Plan 2.

**Procedure for control group (C)**

On the day the second lesson plan assignment was given in class, the following procedures were followed for the control group (C).

**Lesson Plan #2, Day 1 (Introduction)**

1. Participants completed the self-efficacy (pre-task) instrument and returned it to the envelope.

2. Participants sealed their envelope before it was collected.

3. Instructor presented his or her own second lesson plan assignment, as planned.
4. Participants completed Lesson Plan 2 outside class time and submitted it at a later time, typically 2-3 weeks later.

**Lesson Plan #2, Day 2 (Finish)**

1. Participants submitted Lesson Plan 2 on the due date to their course instructor.
2. Participants received envelope containing the self-efficacy posttask and survey.
3. Participants first completed self-efficacy instrument and returned it to the envelope.
4. Participants then removed and completed the questionnaire, before returning it to the envelope.
5. Participants sealed their envelope.
6. Envelopes collected.
7. Later, copies of participants’ lesson plans were provided to the researcher for a separate evaluation process, which did not affect participant’s grade in course.

**Procedure for rubric without explanation group (R)**

On the day the second lesson plan assignment was given in class, the following procedures were followed for the Rubric Group (R).

**Lesson Plan #2, Day 1 (Introduction)**

1. Participant completed the self-efficacy (pre-task) instrument and returned it to the envelope.
2. Participants sealed their envelope before it was collected.

3. Instructor presented his or her own second lesson plan assignment as planned.

4. Participants received the rubric for the lesson plan assignment and were informed their assignment would be evaluated using the rubric.

5. Participants completed their Lesson Plan 2 outside class time and submitted it at a later time.

**Lesson Plan #2, Day 2 (Finish)**

1. Participants submitted Lesson Plan 2 on the due date to their course instructor.

2. Participants received an envelope containing the self-efficacy posttask and survey.

3. Participants completed the self-efficacy instrument and returned it to the envelope.

4. Participants removed and completed the questionnaire, before returning it to the envelope.

5. Participants sealed their envelopes.

6. Envelopes collected.

7. Later, copies of participants’ lesson plans were provided to the researcher for a separate evaluation process, which did not affect participant’s grade in course.
Procedure for rubric with explanation group (RE)

On the day the second lesson plan assignment was given in class, the following procedures were followed for the rubric with explanation group (RE).

Lesson Plan #2, Day 1 (Introduction)

1. Participant completed the self-efficacy (pre-task) instrument and returned it to the envelope.

2. Participants sealed their envelope before it was collected.

3. Instructor presented his or her own second lesson plan assignment as planned.

4. Participants listened to a brief scripted PowerPoint presentation:
   a. on the value of the rubric for increasing academic achievement
   b. using the rubric as a self-assessment, feedback process
   c. encouraging participants to do their best to create the lesson plan according to the criteria in the researcher’s rubric
   d. stressing that their performance on the lesson plan will be determined by the rubric

5. Participants received the rubric for evaluating the lesson plan assignment and were informed the rubric will used to score Lesson Plan 2.

6. Participants completed their second lesson plan outside class time and submitted it at a later time.
Lesson Plan #2, Day 2 (Finish)

1. Participants submitted Lesson Plan 2 on the due date to their course instructor.
2. Participants received an envelope containing the self-efficacy posttask and survey.
3. Participants completed the self-efficacy instrument and returned it to the envelope.
4. Participants removed and completed the questionnaire, before returning it to the envelope.
5. Participants sealed their envelopes.
6. Envelopes were collected.
7. Later, copies of participants’ lesson plans were provided to the researcher for a separate evaluation process, which did not affect participant’s grade in course.

Occurrences That May Have Affected the Procedures

Implementation of the study was hindered in two significant areas: the researcher was not permitted to teach participants in the rubric with explanation group classrooms and the logistical challenge of being present in every classroom for every step in the procedure. Requirements of the state in which the study took place specify that instructors of teacher credential candidates have certification, which the researcher did not. For that reason the course instructors delivered the instructional intervention to the participants. The second challenge was that the number, location, and timing of the research events among multiple courses in one semester logistically prevented the researcher from being present to observe each and every event. To eliminate the bias
that would have occurred if the researcher handled some events for one course, but not for another or for all, it was decided that the course instructors should handle all research events in close collaboration with the researcher.

Changes in the course schedule made by course instructors occurred in a manner typical to any course. One consequence of schedule changes included variation in the amount of time given for participants in some courses to complete the lesson plan assignment. Instructors A, C, and D reported that the sequence and timing in which a few participants completed lesson plan 1, received feedback, and then completed lesson plan 2, differed from the prescribed procedure. For example, in a few cases, a participant did not complete lesson plan 1 until after lesson plan 2 was begun. In another occurrence, the start date for lesson plan 2 in all EDEL 143a courses (Instructors A, B, C, D, E), varied according to the date of the implementation of lesson plan 1. In all courses, the administration of the posttask self-efficacy and survey instruments occurred on a date after the last participant completed lesson plan 2. However, course Instructors did not report any glitches during the actual administration of the instruments.

**Measures to Establish Controls**

In order to reduce errors in the study procedure, course instructors were briefed on the procedure beforehand and given a job aid with the detailed procedure. In addition, each course instructor’s progress through the study procedure was monitored and just prior to beginning a step that required another set of instruments or materials, another briefing occurred to review the next procedural steps.
The most notable flaw in the procedure is the lack of researcher control of the administration of the study in each of the nine classes, requiring heavy dependence upon the research experience and ability of the course instructors.

**Research Design**

The study was a quasi-experimental nonequivalent control group design similar to the nonequivalent control group design described by Campbell and Stanley (1963). The primary purpose of this study was to investigate the effect of three strategies for implementation of rubrics—no rubric (C), rubric given without explanation (R), and rubric given with explanation (RE)—on teacher credential candidates’ self-efficacy and achievement for writing a lesson plan.

Pre-task self-efficacy (PreSE) and Lesson Plan 1 (LP1) were used to determine equivalency among groups before treatment. Post-task self-efficacy (PostSE) was measured to assess the degree of change after the participant had completed the treatment and second lesson plan (LP2). LP2 was the dependent variable for academic achievement.

Participants were enrolled in nine intact classes. Those classes were assigned to one of the three treatment groups. Assignments were made based on each course instructor’s condition for participation in the study.

**Validity**

Because of how the courses were formed by the registration processes, enrollment options, and participants’ preferences for class times or locations, it was believed that participants would be academically equivalent with similar personal characteristics or beliefs, like self-efficacy, and would also be randomly distributed in
the course sections. In order to assess equivalency among groups, PreSE and LP1 were examined.

Because the participants in the intact courses were not randomly assigned to conditions, differences found on LP1 may have been due to preexisting group differences rather than the treatments. To understand the extent to which this may have been true, PreSE and survey items related to preexisting conditions were analyzed to determine whether other significant differences existed. LP1 was used a covariate to analyze any possible differences on LP2.

**Threat to internal validity**

According to Campbell and Stanley (1963) the effectiveness of the nonequivalent control group design increases the more similar the experimental and control groups are in their recruitment and even more so when the similarity is confirmed by the scores on the pretest. Participants for this study were recruited from comparable classes and given a self-efficacy pretest (PreSE) instrument for score comparison. The design controls the main effects of maturation, testing, and instrumentation, but not history. The time lapse between receiving the treatment and actually creating the lesson plan was not controlled in this study. While history may have been a source for extraneous variance, the research design closely matched actual educational practice. Similarly, participants completed the research task of writing a lesson plan on their own, using whatever resources they had available without guarantee of working independently or within the task guidelines, which also matched educational practice. And, no significant events were noted which may have influenced lesson planning skill development or self-efficacy toward lesson planning.
Threat to external validity

The results of this study can be generalized from the sample of students preparing to become teachers (teacher credential candidates) to the same population at comparable higher education institutions. The interaction of testing for self-efficacy and subsequent task achievement was a weakness of this design. The unique nature of the self-efficacy pretest was likely to sensitize participants to their level of confidence for accomplishing the task of creating a lesson plan. The interaction effects of selection bias was minimized by using intact classes, where every student had the opportunity to be a participant in the study and none were selected by any particular criteria. This arrangement reduced the participants’ awareness of participating in a study. Reactive effects of the experimental arrangements were a minimal threat to external validity. The context in which participants were asked to perform the lesson writing task in this study did not differ from the natural setting in which they might have ordinarily performed because writing the lesson plan was a normal assignment in the course.

Data Analysis

The MiniTab version 12 for Windows software program was used to conduct most statistical analysis. SPSS version 11 was used to compute Cronbach’s Alpha reliability coefficient. The following describe the data analysis procedures for the two hypotheses and the research questions.

For the analysis of Hypotheses 1 and 2, reliability-corrected ANCOVA (analysis of covariance) was used to handle measurement error introduced by nonrandom assignment of participants to treatment groups (Tabachnick & Fidell, 2006). In this nonequivalent control group design, the pretest measures (PreSE and LP1) were used as covariates for the posttest measure (PostSE and LP2, respectively) to reduce the amount
of “noise” caused by the measurement error (Trochim, 2006). Data on the covariates was gathered before treatment in order to meet the condition that the covariates were independent of treatments.

The rationale for using reliability-corrected ANCOVA is that a standard ANCOVA analysis is biased when used with a nonequivalent group design. The bias is introduced by a combination of pretest measurement error and group nonequivalence (Trochim, 2006). Since the problem is partly due to measurement error on the pretest, the solution is to remove the pretest measurement error and approximate the no pretest error case. Trochim (2006) further explains that the “reliability” correction adjusts the pretest for measurement error by reducing the amount of error, thus squeezing the pretest distribution inwards proportionate to the amount that measurement error stretched it. Using Cronbach’s Alpha reliability coefficient, the formula for the adjustment is:

$$X_{adj} = \bar{X} + r(X - \bar{X})$$

where:

- $X_{adj}$ = adjusted pretest value
- $\bar{X}$ = original mean pretest value
- $r$ = Cronbach’s Alpha reliability coefficient

Given reliability is an estimate of the proportion of the measure that is true score relative to error, it becomes the factor by which the raw scores must be adjusted. The premise was to construct a new set of pretest scores for each participant for PreSE and LP1. These new scores reflected the adjustment for pretest unreliability by an amount in proportion to the reliability. In effect, a participant’s adjusted pretest score was closer to the mean for that group. The same reliability adjustment method was used to account for the unreliability in the post-treatment measures, PostSE and LP1.
Cronbach’s Alpha was chosen because it is a frequently used estimate of internal consistency and uses a robust computation algorithm to estimate reliability (Trochim, 2006).

For analysis of the research questions, the chi square test was used to analyze the categorical data from the survey to determine statistical significance among relationships observed in the contingency tables (Gall, Gall, & Borg, 2005).
The hypotheses and research questions, data collection methods, instruments, and data analysis methods are presented in Table 3.6 and Table 3.7.

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Data Collection Methods</th>
<th>Instruments</th>
<th>Data Analysis Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₀₁</td>
<td>Research study</td>
<td>Self-efficacy instrument</td>
<td>ANCOVA</td>
</tr>
<tr>
<td>There will be no significant difference in the means for post-task self-efficacy for writing a lesson plan among three treatment groups, no rubric, rubric given without explanation, and rubric given with explanation.</td>
<td>post-task self-efficacy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| H₀₂        | Research study          | Lesson plan 2 (LP2) | ANCOVA               |
| There will be no significant difference in the means for achievement for writing a lesson plan among three treatment groups, no rubric, rubric given without explanation, and rubric given with explanation. | lesson plans |                     |

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Data Collection Methods</th>
<th>Instruments</th>
<th>Data Analysis Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1.</td>
<td>Research study</td>
<td>Self-reported post-survey</td>
<td>Chi square</td>
</tr>
<tr>
<td>Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in the amount of time spent creating the second lesson plan?</td>
<td>survey</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Q2.                | Research study          | Self-reported post-survey | Chi square          |
| Do the two strategies for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of the usefulness of the rubric? | survey |                     |
Table 3.7  Research Questions Based on the Survey, Data Collection Methods, Instruments, and Data Analysis Methods

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Data Collection Methods</th>
<th>Instruments</th>
<th>Data Analysis Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q3a. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of how to write a lesson plan before starting the task?</td>
<td>Research study</td>
<td>Self-reported post-survey</td>
<td>Chi square</td>
</tr>
<tr>
<td></td>
<td>research study</td>
<td>survey</td>
<td></td>
</tr>
<tr>
<td>Q3b. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ understanding of the requirements to earn the highest possible score on the lesson plan assignment?</td>
<td>Research study</td>
<td>Self-reported post-survey</td>
<td>Chi square</td>
</tr>
<tr>
<td></td>
<td>research study</td>
<td>survey</td>
<td></td>
</tr>
<tr>
<td>Q3c. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of the helpfulness of the rubric for setting goals to create the lesson plan?</td>
<td>Research study</td>
<td>Self-reported post-survey</td>
<td>Chi square</td>
</tr>
<tr>
<td></td>
<td>research study</td>
<td>survey</td>
<td></td>
</tr>
<tr>
<td>Q3d. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of understandable feedback derived from using the rubric?</td>
<td>Research study</td>
<td>Self-reported post-survey</td>
<td>Chi square</td>
</tr>
<tr>
<td></td>
<td>research study</td>
<td>survey</td>
<td></td>
</tr>
<tr>
<td>Q3e. Does the strategy for implementation of analytic trait, task-specific rubric use lead to different amounts of time saved completing the lesson plan writing assignment?</td>
<td>Research study</td>
<td>Self-reported post-survey</td>
<td>Chi square</td>
</tr>
<tr>
<td></td>
<td>research study</td>
<td>survey</td>
<td></td>
</tr>
<tr>
<td>Q3f. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of using the rubric for self-evaluation while creating the lesson plan?</td>
<td>Research study</td>
<td>Self-reported post-survey</td>
<td>Chi square</td>
</tr>
<tr>
<td></td>
<td>research study</td>
<td>survey</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 4
RESULTS

This chapter presents the results of the study. The primary purpose of this quasi-experimental study was to investigate the effect of three strategies for implementation of an analytic trait, task-specific rubric—no rubric (C), rubric given without explanation (R), and rubric given with explanation (RE)—on teacher credential candidates’ self-efficacy and achievement for writing a lesson plan. Specifically, the research questions were:

1. Which implementation of analytic trait, task-specific rubrics is more effective in increasing self-efficacy for writing a lesson plan?
2. Which implementation of analytic trait, task-specific rubrics is more effective in increasing achievement for writing a lesson plan?

The study was implemented in the Fall of 2005 and Spring of 2006 at a large, west coast, public university with 114 volunteer participants recruited from courses for teacher credential candidates, where the typical student in the course had little or no prior experience writing lesson plans. The participants were in intact courses and not randomly assigned to treatments. In quasi- or nonexperimental situations like this case, ANCOVA is used as a statistical matching procedure to adjusting the group mean scores (Tabachnick & Fidell, 2007) to control for the influence of self-efficacy before treatment. This reliability-corrected ANCOVA model is used in the statistical analysis of the nonequivalent control group design to correct for the bias that would occur as a result of measurement error on the pretest (Trochim, 2006).

Data analysis results are presented in two major parts; first by the conditions prior to treatment, which includes the data regarding equivalency, the pre-treatment
self-efficacy, and lesson plan 1 and second by the conditions after treatment, including post-treatment self-efficacy, lesson plan 2, and responses to survey items regarding participants’ perceptions of the study. The analysis of the survey data is presented within each of the two parts based on whether a survey item was related to a pre-existing condition or a post-treatment condition. In addition, correlations examining relationships among the dependent variables and among survey items are shown. At the end of the chapter is a summary of the findings relating to the null hypotheses tested in this study.

**Summary of Hypotheses**

**Hypothesis 1.** There will be no significant difference between the means for post-treatment self-efficacy among three treatment groups.

The above null hypothesis was rejected at the .05 level, because the control group (C) exhibited significantly higher self-efficacy after the treatment.

**Hypothesis 2.** There will be no significant difference between the means for achievement for writing a lesson plan among three treatment groups, no rubric, rubric given without explanation, and rubric given with explanation.

The above null hypothesis was retained at the .05 level.

**Results: Pre-Treatment Conditions**

The data in this section are presented to create a demographic profile of the participants, and then to determine the equivalency of groups before treatment in order to better interpret whether any differences found in the post-treatment measures were due to the treatment or other factors.
Equivalency of Settings

The settings in which courses were held varied as shown in Table 4.1. Course instructors had conducted their classes in these settings before and reported no detriment to the study. The impact of the different learning environments on the study remains unclear.
Table 4.1  Course Settings

<table>
<thead>
<tr>
<th>Sem</th>
<th>Instr</th>
<th>Course</th>
<th>Location</th>
<th>Room Type</th>
<th>Envrnmnt</th>
<th>Day / Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2005</td>
<td>A</td>
<td>EDEL</td>
<td>Off-campus</td>
<td>high school classroom</td>
<td>seminar style</td>
<td>Friday AM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>143a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring 2006</td>
<td>B</td>
<td>EDEL</td>
<td>Off-campus</td>
<td>restaurants and elementary</td>
<td>seminar style</td>
<td>Tues AM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>143a</td>
<td></td>
<td>school meeting rooms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>C</td>
<td>EDEL</td>
<td>Off-campus</td>
<td>school district office meeting room</td>
<td>large-group lecture and discussion</td>
<td>Friday AM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>143a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>D</td>
<td>EDEL</td>
<td>Off-campus</td>
<td>elementary school library, seminar style</td>
<td></td>
<td>Thurs PM (eve)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>143a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>E</td>
<td>EDEL</td>
<td>Off-campus</td>
<td>elementary school district office</td>
<td>large-group lecture and discussion</td>
<td>Wed AM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>143a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>F</td>
<td>ChAD</td>
<td>On-campus</td>
<td>large lecture hall with fixed seats</td>
<td>lecture and discussion</td>
<td>Mon &amp; Wed PM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>149</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>G</td>
<td>EDIT</td>
<td>On-campus</td>
<td>computer lab</td>
<td>lecture and discussion</td>
<td>Mon PM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>122</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>G</td>
<td>EDIT</td>
<td>On-campus</td>
<td>computer lab</td>
<td>lecture and discussion</td>
<td>Tues PM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>241</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring 2006</td>
<td>G</td>
<td>EDIT</td>
<td>On-campus</td>
<td>computer lab</td>
<td>lecture and discussion</td>
<td>Wed PM (eve)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>241</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Equivalency of Course Instructors

Each course instructor taught in a different style and emphasized different aspects of lesson plan development, as would be expected of faculty in a large university. Each instructor had taught the course before the study and had expressed confidence they could implement the study into their regular curriculum without difficulty. The five instructors who taught EDEL 143a regularly collaborated with all faculty who taught that course to confirm a common set of instructional outcomes for the teacher credential candidates. Equivalency among EDEL 143a instructors was expected to be higher among themselves than the other instructors.

During the treatment, differences among instructors style of teaching became more evident, such as emphasis on the details of lesson plan writing, the timing of feedback on lesson plan performance, the duration of the two lesson plan assignments, and the sequencing of the lesson plan assignment within the course curriculum and schedule. In addition, it became evident from debriefing conversations with instructors that each conveyed their own emphasis on the research materials and instruments used in the study to the participants, which varied from the intended emphasis and tone of the study. In particular, differences were noted in the degree of seriousness with which participants were informed to approach the study. While all data was handled in accordance with established institutional review board and office of research protections policy, the procedures of the study, as described in the research procedure, were not consistently implemented across the instructors.

A few course instructors responded to an email inquiry about the outcomes of the study and any unusual occurrences. Instructor B (control group) reported no unusual occurrences during the study. Instructor C (rubric without explanation) reported the following unusual conditions that occurred during the study period.
I was sharing [co-teaching] my instruction with a colleague, and we had difficulty collaborating. We had differing attitudes about our expectations of the students, about the content of the course and about doing the study and I’m sure the students picked up on this, although nothing specific was ever expressed.

In addition, I must say that the group of students in the Fall 2005 cohort (the study group) was negatively affected by a small number of individuals with strong personalities who were very vocal.

Finally, in order to respond to complaints of being overworked, we tried to combine assignments so that their coaching cycle lesson plan, which fulfilled a requirement for their language arts class, met the requirement for their 143a lesson plan; therefore, we felt inhibited about providing feedback in how to teach the strategies, so our comments were very general, not my typical style. The coaching cycle document [Appendix G] is what I used...I find their written reflections revealing because they are required to explain their thinking.

**Participant demographics**

Table 4.2 reports the assignment of nine intact classes to three treatment groups: C is the control (n=39), R is the rubric without explanation (n=36), and RE is rubric with explanation (n=39). Assignments were made based on each course instructor’s condition for participation in the study. Those instructors who expressed commitment to presenting the 20-minute presentation on rubric use were assigned to RE. Instructors who were willing to give the rubric to participants, but not the presentation, were assigned to R. Instructors who were willing to give only the pre- and post-task self-efficacy instrument, were assigned to C. Treatment groups of at least forty participants were arranged based on these conditions of instructor participation and number of enrollments per class. Attrition occurred during the study that reduced the final participant count for each group.

The majority of participants in each course were teacher credential candidates in the early phase of their academic program. Each was required to write lesson plans as part of their normal course work. Participants typically selected their courses based on a
combination of preferences for schedule, instructor, or specialized credential program. Elementary Education refers to five sections of the course EDEL 143a, “Orientation to Student Teaching” and was represented in each treatment group by at least 30% of the participants. Child and Adolescent Development (ChAD) refers to ChAD 149, “Child Health and Physical Activity,” which comprised nearly 70% of the participants in control (C) group. “Microcomputers in Education,” (EDIT 122) was an entry-level technology course for teachers and represented 11% of the participants in the R group. The two sections of “Emerging Technologies for Teachers,” (EDIT 241) were the second-level technology course for teachers and represented 39% of the R group and nearly 18% of the RE group.

Comparing treatment groups, the data shows one-third of the participants in each group were from one or two of the Elementary Education courses. The control group had the addition of participants from ChAD, while R was comprised of participants in two different technology integration courses, EDIT 122 and EDIT 241, and RE had only seven additional participants from one technology integration course.
Table 4.2  Participants by Course and Treatment (n=114)

<table>
<thead>
<tr>
<th>Semester</th>
<th>Instr.</th>
<th>Course</th>
<th>Control (C)</th>
<th>%</th>
<th>Rubric without explanation (R)</th>
<th>%</th>
<th>Rubric with explanation (RE)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2005</td>
<td>A</td>
<td>EDEL 143a</td>
<td>4</td>
<td>10.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring 2006</td>
<td>B</td>
<td>EDEL 143a</td>
<td>8</td>
<td>20.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>C</td>
<td>EDEL 143a</td>
<td></td>
<td></td>
<td>18</td>
<td>50.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>D</td>
<td>EDEL 143a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>38.46</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>E</td>
<td>EDEL 143a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
<td>43.59</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>F</td>
<td>ChAD 149</td>
<td>27</td>
<td>69.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall 2005</td>
<td>G</td>
<td>EDIT 122</td>
<td>4</td>
<td>11.0</td>
<td></td>
<td></td>
<td>4</td>
<td>11.0</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>G</td>
<td>EDIT 241</td>
<td>14</td>
<td>39.0</td>
<td></td>
<td></td>
<td>14</td>
<td>39.0</td>
</tr>
<tr>
<td>Spring 2006</td>
<td>G</td>
<td>EDIT 241</td>
<td>7</td>
<td>17.95</td>
<td></td>
<td></td>
<td>7</td>
<td>17.95</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>39</td>
<td>100.0</td>
<td>36</td>
<td>100.0</td>
<td>39</td>
<td>100.0</td>
</tr>
</tbody>
</table>

**Participant majors and academic programs**

As shown in Table 4.3, more than 78% of total responses to survey item 1 indicated the participants’ major or academic program was in the teacher credential program in the College of Education. The remaining participants were classified as Liberal Studies (9.82%), Other Education (9.82%), and Other (11.61%). “Other” was comprised of participants whose major fields of study were chemistry, creative arts, economics, English, history, physics, or psychology, and were enrolled in teacher preparation courses. The category named “Other Education,” classified teacher credential candidates who were in advanced stages of teacher preparation or early teaching service.
The control group (C) was comprised of a majority (54%) of teacher credential candidates in the following majors: ChAD, Elementary Education, Multiple Subjects, and Teaching Credential. The remainder was in Liberal Studies and Other. For group R, nearly 92% of participants indicated their major was in a teacher credential program as were nearly 90% for group RE. The most diverse group by major was RE, which comprised all represented majors except Critical Research Academy. The majority of RE participants (78.57%) were in the teacher credential program with the remainder in Liberal Studies and Other. Overall, the total percentage of majors represented among all participants shows a fairly equal distribution in the range of 5% - 13% with the exception of the high percentage of participants from Multiple Subjects (18.75%).
### Table 4.3 Academic Major by Treatment

<table>
<thead>
<tr>
<th>Major</th>
<th>Control (C)</th>
<th>Rubric (R)</th>
<th>Rubric+ (RE)</th>
<th>Total by Major</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>Percent</td>
<td>Count</td>
<td>Percent</td>
</tr>
<tr>
<td>BCLAD</td>
<td>10</td>
<td>25.64</td>
<td>1</td>
<td>2.78</td>
</tr>
<tr>
<td>ChAD (Child &amp; Adolescent Ed)</td>
<td></td>
<td></td>
<td>14</td>
<td>38.89</td>
</tr>
<tr>
<td>Elementary Ed</td>
<td>3</td>
<td>7.69</td>
<td>1</td>
<td>2.78</td>
</tr>
<tr>
<td>Liberal Studies</td>
<td>10</td>
<td>25.64</td>
<td>1</td>
<td>2.70</td>
</tr>
<tr>
<td>Multiple Subjects</td>
<td>6</td>
<td>15.38</td>
<td>5</td>
<td>13.89</td>
</tr>
<tr>
<td>SUP Internship</td>
<td></td>
<td></td>
<td>6</td>
<td>16.22</td>
</tr>
<tr>
<td>Teaching Credential</td>
<td>2</td>
<td>5.13</td>
<td>5</td>
<td>13.89</td>
</tr>
<tr>
<td>Other Education</td>
<td>7</td>
<td>19.44</td>
<td>4</td>
<td>10.81</td>
</tr>
<tr>
<td>Other</td>
<td>8</td>
<td>20.51</td>
<td>3</td>
<td>8.33</td>
</tr>
<tr>
<td>Total (N)</td>
<td>39</td>
<td>100.00</td>
<td>36</td>
<td>100.00</td>
</tr>
</tbody>
</table>

### Summary of Equivalency of Treatment Groups

Given the condition of this study that participants were not randomly assigned to treatment groups and that groups were not randomly assigned to treatments, the problem of combined pretest measurement error and group nonequivalence must be
addressed in order to correctly apply ANCOVA on the post-treatment measures, PostSE and LP2. Measurement error can be corrected with a formula built around Cronbach’s Alpha reliability coefficient \((r)\) for the pre-treatment measures, PreSE and LP1. Trochim (2006) describes the formula for adjustment:

\[
X_{\text{adj}} = \bar{X} + r \left( X - \bar{X} \right).
\]

Group nonequivalence can be better understood in the context of the data that describes the conditions of the study before treatment was administered. The preceding section described the key conditions that may have an impact on the results. In summary, the settings of each course varied from traditional classroom to seminar room to restaurant, but the impact of the different settings on this study is not clear because students in the past had been able to competently perform within these environments. The course instructors were an important factor in the implementation of this study and extreme differences were discovered in their teaching style, the lesson plan assignment they gave to participants, and how they managed the study. Via email survey and conversations with the course instructors, the differences became known and were likely to have impact on the results of the study. In particular, instructors offered different types and forms of feedback to the participants was likely to have influenced them when completing the lesson plan assignments. In this naturalistic setting it was expected that instructors controlled and directed the participants regarding the requirements of the study and response to the instruments or intervention, just like any other assignment in the course. Differences among participants were most notable in terms of course selection, which in part, was dependent on self-selection for their particular credential program. The differences in course content influenced the nature of the lesson plan assignment and corresponding requirements for writing the lesson.
The matched assignment of classes to treatments created unbalanced groups of participants with respect to clusters of certain majors being in a treatment. A participant’s major is a self-selection based on many factors including personal interest, previous experience, skill, knowledge, and schedule availability.

Taken together these factors introduce multiple variables that were not controlled in the naturalistic setting of the study. The remaining results should be interpreted with caution and not generalized outside the scope of this study.

**Pre-treatment Self-efficacy**

A self-efficacy instrument (Appendix B) was designed for this study to measure participants’ levels of confidence for specific abilities and subtasks pertinent to writing a lesson plan and for overall confidence in their academic achievement for the task. The two-page instrument contained a total of sixteen items. High internal reliability was measured using Cronbach’s Alpha ($r=.95$). The first item, set alone on the first page, asked the participants to identify their level of confidence for getting an “A” grade on the lesson plan assignment. The second page contained the remaining items.

**Overall confidence to achieve an “A” grade**

The first item on the instrument asked participants to circle the number from 1 “not sure” to 10 “very sure” that best described the truth of the following statement: “I am confident I will get an ‘A’ grade on the next lesson plan assignment in this course.” Table 4.4 contains the descriptive statistics. The overall mean ($M = 7.13$) corresponds to the description, “pretty sure,” that the participant would get an “A” grade on the next lesson plan, which would be the lesson plan completed as part of the treatment in this study.
Table 4.4  Descriptive Data for Level of Confidence to Get an “A” grade on Lesson Plan Assignment

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>39</td>
<td>6.95</td>
<td>2.34</td>
<td>0.38</td>
<td>1.00</td>
<td>10.00</td>
</tr>
<tr>
<td>R</td>
<td>36</td>
<td>7.53</td>
<td>2.21</td>
<td>0.37</td>
<td>2.00</td>
<td>10.00</td>
</tr>
<tr>
<td>RE</td>
<td>39</td>
<td>6.95</td>
<td>1.89</td>
<td>0.30</td>
<td>3.00</td>
<td>10.00</td>
</tr>
<tr>
<td>Total</td>
<td>114</td>
<td>7.13</td>
<td>2.17</td>
<td>0.20</td>
<td>1.00</td>
<td>10.00</td>
</tr>
</tbody>
</table>

Self-efficacy for writing a lesson plan

The level of self-efficacy for writing a lesson plan before the treatment was assessed using the overall mean score obtained from 14 items on the instrument. Each item had a scale with a minimum score of 1 “not sure” and a maximum score of 10 “very sure.” Overall mean scores of the 14 items in the range 1-3 indicated low levels of self-efficacy, a mean of 4-6 indicated a middle level, and 7-10 indicated a high level self-efficacy. Table 4.5, reports the adjusted descriptive data for self-efficacy before treatment (PreSE). To account for the nonrandom assignment of participants to treatments, the mean scores were adjusted to partly account for measurement error on PreSE using a "reliability" correction that adjusts the pretest for measurement error. The formula for adjustment squeezes the PreSE distribution inwards to correctly adjust for measurement error. Adjusted PreSE helps to control the bias in ANCOVA treatment effect estimates for the nonequivalent control group design.
Table 4.5  Descriptive Data for Self-efficacy Before Treatment (PreSE)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>39</td>
<td>7.77</td>
<td>1.44</td>
<td>.23</td>
<td>3.42</td>
<td>9.66</td>
</tr>
<tr>
<td>R</td>
<td>36</td>
<td>8.04</td>
<td>1.11</td>
<td>.19</td>
<td>5.69</td>
<td>9.89</td>
</tr>
<tr>
<td>RE</td>
<td>39</td>
<td>7.81</td>
<td>1.33</td>
<td>.21</td>
<td>4.85</td>
<td>9.89</td>
</tr>
<tr>
<td>Total</td>
<td>114</td>
<td>7.87</td>
<td>1.02</td>
<td>.13</td>
<td>3.42</td>
<td>9.89</td>
</tr>
</tbody>
</table>

The descriptive data shows that the adjusted scores ranged from 3.42 to 9.89 among all participants. R group showed the smallest range, 5.69 to 9.89, with the lowest standard deviation (1.11) and the highest mean (M = 8.04) among all treatment groups. The means for all treatment groups lay within the high range (7-10) for self-efficacy. Analysis of PostSE using reliability-corrected ANCOVA to partial out the effect of PreSE is presented later in this chapter.

**Pre-treatment Achievement: Lesson Plan 1**

Achievement was calculated by the average of two raters’ scores on the second lesson plan (LP2) after treatment. Two independent raters in a blind-review process determined the scores.

**Inter-rater reliability of lesson plan scores**

Table 4.6 shows the correlation between the ratings of each pair of raters for LP1 and LP2. The conservative interpretation of correlations describes values in the range greater than or equal to .70 to less than .90, as high correlation and values in the range greater than .50 to less than .70, as moderate correlation.
Table 4.6  Inter-rater Reliability Correlations

<table>
<thead>
<tr>
<th>Rater 1 Code</th>
<th>Rater 2 Code</th>
<th>Semester</th>
<th>Correlation (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM*</td>
<td>KF</td>
<td>Fall 2005</td>
<td>0.69</td>
</tr>
<tr>
<td>MH</td>
<td>KH</td>
<td>Fall 2005</td>
<td>0.68</td>
</tr>
<tr>
<td>ML*</td>
<td>VP</td>
<td>Fall 2005</td>
<td>0.82</td>
</tr>
<tr>
<td>MS1</td>
<td>MS2</td>
<td>Fall 2005</td>
<td>0.77</td>
</tr>
<tr>
<td>UF*</td>
<td>SC</td>
<td>Fall 2005</td>
<td>0.66</td>
</tr>
<tr>
<td>CS</td>
<td>UF</td>
<td>Spring 2006</td>
<td>0.54</td>
</tr>
<tr>
<td>CO</td>
<td>KM</td>
<td>Spring 2006</td>
<td>0.70</td>
</tr>
<tr>
<td>ML</td>
<td>AM</td>
<td>Spring 2006</td>
<td>0.69</td>
</tr>
</tbody>
</table>

*Average 0.69
*rater participated twice

Lesson plan 1

Table 4.7 reports the adjusted descriptive data for lesson plan 1 (LP1).

Table 4.7  Adjusted Descriptive Data for Achievement on Lesson Plan 1 (LP1)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>39</td>
<td>26.11</td>
<td>8.33</td>
<td>1.33</td>
<td>6.87</td>
<td>43.76</td>
</tr>
<tr>
<td>R</td>
<td>36</td>
<td>31.28</td>
<td>5.99</td>
<td>1.0</td>
<td>15.58</td>
<td>41.69</td>
</tr>
<tr>
<td>RE</td>
<td>39</td>
<td>27.04</td>
<td>6.94</td>
<td>1.11</td>
<td>7.29</td>
<td>42.94</td>
</tr>
<tr>
<td>Total</td>
<td>114</td>
<td>28.06</td>
<td>7.46</td>
<td>.843</td>
<td>6.87</td>
<td>43.76</td>
</tr>
</tbody>
</table>

The adjusted descriptive data above shows that the achievement on LP1 ranged widely from 6.87 to 43.76 out of 51 for the C group, from 15.58 to 41.69 for the R group, and 7.29 to 42.94 for the RE group. Overall, the mean was 28.06 with a standard deviation of 7.46. Analysis of LP2 using reliability-corrected ANCOVA to partial out the effect of LP1 is presented later in this chapter.
Survey Results: Pre Treatment Items

To further investigate equivalency among treatment groups, data about participants were also obtained from a 15-item survey that was administered after the treatment (see Appendix A). An overview of the survey instrument and the results of items 2, 3, 5, 6, and 7, which pertain to conditions before treatment, are presented in this section. The results of the remaining items (4, 8, and 9a-f) that may have been affected by the treatment are presented in the following section.

Every participant was given the survey after they had received the treatment and completed the second lesson plan assignment, LP2. Each was instructed to answer items 1–7, and in addition, treatment groups R and RE were instructed to answer items 8 and 9a-f, which pertained to rubric use. To prevent patterned responses, the response choices for items 5 and 6 were reversed so the first choice options, “always” and “frequently” respectively, were the highest level of frequency response options for the item.

Prior Teaching Experience

Table 4.8 shows participants’ response to survey item 2, “...years that matches your total classroom teaching experience.” Before chi-square tests were conducted, each participant’s response for item 2 was coded. The six response options were collapsed into three categories, “Less than 1 year,” “1-3 years,” and “More than 3 years.” A chi-square test of independence was calculated to test the probability of a relationship among treatment groups and years of teaching experience.
Table 4.8   Years of Teaching Experience by Treatment

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Count</th>
<th>% of group</th>
<th>% of item</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Control (C)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=37)</td>
<td>27</td>
<td>73.0%</td>
<td>43.6%</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>18.9%</td>
<td>21.2%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>8.1%</td>
<td>18.8%</td>
</tr>
<tr>
<td><strong>Rubric (R)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=36)</td>
<td>17</td>
<td>47.2%</td>
<td>27.4%</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>33.3%</td>
<td>36.4%</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>19.4%</td>
<td>43.8%</td>
</tr>
<tr>
<td><strong>Rubric+ (RE)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=38)</td>
<td>18</td>
<td>47.4%</td>
<td>29.0%</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>36.8%</td>
<td>42.4%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>15.8%</td>
<td>37.5%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>62</td>
<td>55.9%</td>
<td>29.7%</td>
</tr>
<tr>
<td>(n=111)</td>
<td>33</td>
<td>29.7%</td>
<td>14.4%</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Chi-Square = 6.872, DF = 4, P-Value = 0.143

Overall, more than half of participants (55.9%) indicated they had less than 1 year of experience and nearly 30% had 1-3 years of experience. The control group had the highest percentage of participants with less than 1 year of experience (43.6%) and R (27.4%) and RE (29%) were nearly equal. The overall percentage of participants with more than 3 years experience was 14.4%. The result of the chi-square test of independence indicated there was not a significant relationship among years of teaching experience and treatment ($\chi^2 = 6.872, df = 4, p = 0.143$). The non-significant
relationship could be expected because the participants were enrolled in introductory teacher preparation courses.

To further determine participants’ prior teaching experience with respect to a particular skill required of teachers, Table 4.9, shows results of survey item 3, “…level of experience creating lesson plans.” The four response choices on the survey were collapsed into three categories, “very little,” “some,” and “quite a bit.” Because of the low frequency of the choice “a lot,” it was combined with “quite a bit.” A chi-square test of independence was calculated to test the probability of a relationship among treatment groups and experience creating lesson plans.

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Count</th>
<th>% of group</th>
<th>% of item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (C)</td>
<td>12</td>
<td>30.8%</td>
<td>34.3%</td>
</tr>
<tr>
<td>(n=39)</td>
<td></td>
<td>46.1%</td>
<td>36.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23.0%</td>
<td>30.0%</td>
</tr>
<tr>
<td>Rubric (R)</td>
<td>10</td>
<td>27.8%</td>
<td>28.6%</td>
</tr>
<tr>
<td>(n=36)</td>
<td></td>
<td>36.1%</td>
<td>26.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>36.1%</td>
<td>43.3%</td>
</tr>
<tr>
<td>Rubric+ (RE)</td>
<td>13</td>
<td>33.3%</td>
<td>37.1%</td>
</tr>
<tr>
<td>(n=39)</td>
<td></td>
<td>46.2%</td>
<td>36.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20.5%</td>
<td>26.7%</td>
</tr>
<tr>
<td>Total</td>
<td>35</td>
<td>30.7%</td>
<td>30.7%</td>
</tr>
<tr>
<td>(n=114)</td>
<td></td>
<td>43.0%</td>
<td>26.3%</td>
</tr>
</tbody>
</table>

Chi-Square = 2.733, DF = 4, P-Value = 0.603
The data totals for item choices shows the majority of participants classified their level of experience creating lesson plans as “some” (43.0%) or “very little” (30.7%). Overall, the data appear fairly flat without large differences either within groups or among groups. The result of the chi-square test of independence confirmed there was not a significant relationship among level of experience creating lesson plans and treatment ($\chi^2 = 2.733, df = 4, p = 0.603$).

**Prior Experiences Using Rubrics**

Using rubrics can be considered a skill for teachers. The survey contained two items to determine participants’ use of rubrics outside this study. Item 6 asked for participants’ frequency using rubrics in the past to evaluate their own work and item 7, the frequency of using rubrics when evaluating their current students’ work. Table 4.10 shows the data from survey item 6, “In the past, I have used scoring rubrics to evaluate my own work on assignments or projects.”
Table 4.10  Prior Experience Using Rubrics to Evaluate Own Work

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Count</th>
<th>Frequently</th>
<th>Often</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (C)</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>11</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>(n=39)</td>
<td></td>
<td>13.16</td>
<td>21.05</td>
<td>21.05</td>
<td>28.95</td>
<td>15.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50.00</td>
<td>34.78</td>
<td>25.81</td>
<td>36.67</td>
<td>31.58</td>
</tr>
<tr>
<td>Rubric (R)</td>
<td>4</td>
<td>5</td>
<td>13</td>
<td>8</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>(n=36)</td>
<td></td>
<td>11.11</td>
<td>13.89</td>
<td>36.11</td>
<td>22.22</td>
<td>16.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40.00</td>
<td>21.74</td>
<td>41.94</td>
<td>26.67</td>
<td>31.58</td>
</tr>
<tr>
<td>Rubric+ (RE)</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>(n=39)</td>
<td></td>
<td>2.56</td>
<td>25.64</td>
<td>25.64</td>
<td>28.21</td>
<td>17.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.00</td>
<td>43.48</td>
<td>32.26</td>
<td>36.67</td>
<td>36.84</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>23</td>
<td>31</td>
<td>30</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>(n=114)</td>
<td></td>
<td>8.85</td>
<td>20.35</td>
<td>27.43</td>
<td>26.55</td>
<td>16.81</td>
</tr>
</tbody>
</table>

Chi-Square = 6.058, DF = 8, P-Value = 0.641
3 cells with expected counts less than 5.0

The pattern of responses for all participants and within groups (% of group) closely matches a normal curve. Combined percentages for the least frequent use categories, “rarely” and “never,” yields interesting insight that more than one-third of participants in each group report limited use of rubrics: C equals 44.74%, R equals 38.89%, and RE equals 43.36%. The result of the chi-square test of independence
confirmed there was not a significant relationship among participants’ frequency of using rubrics to evaluate their own work and treatment ($\chi^2 = 6.058$, $df = 8$, $p = 0.641$). The data show evidence that participants had limited use of rubrics to evaluate their own work.

Table 4.11 shows the data from survey item 7, “When you teach, how often do you use a scoring rubric to evaluate or score your students’ academic work/projects?” The survey response choice “always” was collapsed into category “often.”

<table>
<thead>
<tr>
<th>Table 4.11 Experience Using Rubrics to Evaluate Students’ Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 7. Use of Rubrics to Evaluate Students’ Work</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Treatment Group</td>
</tr>
<tr>
<td>Count</td>
</tr>
<tr>
<td>% of group</td>
</tr>
<tr>
<td>% of item</td>
</tr>
<tr>
<td>Control (C) (n=39)</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>20.00</td>
</tr>
<tr>
<td>28.57</td>
</tr>
<tr>
<td>31.43</td>
</tr>
<tr>
<td>20.00</td>
</tr>
<tr>
<td>38.89</td>
</tr>
<tr>
<td>24.39</td>
</tr>
<tr>
<td>35.48</td>
</tr>
<tr>
<td>38.89</td>
</tr>
<tr>
<td>Rubric (R) (n=36)</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>17</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>11.43</td>
</tr>
<tr>
<td>48.57</td>
</tr>
<tr>
<td>28.57</td>
</tr>
<tr>
<td>11.43</td>
</tr>
<tr>
<td>22.22</td>
</tr>
<tr>
<td>41.46</td>
</tr>
<tr>
<td>32.26</td>
</tr>
<tr>
<td>22.22</td>
</tr>
<tr>
<td>Rubric+ (RE) (n=39)</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>18.42</td>
</tr>
<tr>
<td>36.84</td>
</tr>
<tr>
<td>26.32</td>
</tr>
<tr>
<td>18.42</td>
</tr>
<tr>
<td>38.89</td>
</tr>
<tr>
<td>34.15</td>
</tr>
<tr>
<td>32.26</td>
</tr>
<tr>
<td>38.89</td>
</tr>
<tr>
<td>Total (n=114)</td>
</tr>
<tr>
<td>18</td>
</tr>
<tr>
<td>41</td>
</tr>
<tr>
<td>31</td>
</tr>
<tr>
<td>18</td>
</tr>
<tr>
<td>16.67</td>
</tr>
<tr>
<td>37.96</td>
</tr>
<tr>
<td>28.70</td>
</tr>
<tr>
<td>16.67</td>
</tr>
</tbody>
</table>

Chi-Square = 3.789, DF = 6, P-Value = 0.705
For all participants’ response choice the data resembles a normal curve with the highest percent response (37.96%) as “rarely” did participants use a rubric to score their students work. When the within groups data (% of group) for “never” and “rarely” are combined the summed percentage is almost half or more than half of the total responses for each group: C equal 48.57%, R equal 60%, and RE equal 55.26%. Overall, the data do not indicate large differences in response choice among groups. The result of the chi-square test of independence confirmed there was not a significant relationship among participants’ frequency of using rubrics to evaluate their students’ work and treatment ($\chi^2 = 3.789, df = 6, p = 0.705$). The data do show evidence that participants had limited use of rubrics to evaluate their students’ work.

**Results: Post Treatment Conditions**

The reliability-corrected ANCOVA model was employed to test for differences among groups in self-efficacy after treatment (PostSE) while controlling for the influence of self-efficacy before treatment (PreSE) and results of the analysis were used to answer research question 1. The reliability-corrected ANCOVA model was also used to test for differences among groups for achievement on LP2 while controlling for the influence of LP1, to answer research question 2.

**Post Treatment Self-efficacy**

**Overall confidence to achieve an “A” grade**

After completing the second lesson plan (LP2) participants were given the same self-efficacy instrument as before the treatment. In response to the first item, “I am confident I will get an ‘A’ grade on the next lesson plan assignment in this course,” Table 4.12 contains the descriptive statistics. The overall mean ($M = 7.68$) corresponds to
the description, “pretty sure,” that the participant would get an “A” grade on the next
lesson plan, which would be the lesson plan completed as part of the treatment in this
study.

<p>| Table 4.12 | Descriptive Data for Level of Confidence to Get an “A” grade on Lesson Plan Assignment |</p>
<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>39</td>
<td>8.13</td>
<td>1.70</td>
<td>0.27</td>
<td>1.00</td>
<td>10.00</td>
</tr>
<tr>
<td>R</td>
<td>36</td>
<td>7.67</td>
<td>1.81</td>
<td>0.30</td>
<td>2.00</td>
<td>10.00</td>
</tr>
<tr>
<td>RE</td>
<td>39</td>
<td>7.23</td>
<td>2.28</td>
<td>0.37</td>
<td>1.00</td>
<td>10.00</td>
</tr>
<tr>
<td>Total</td>
<td>114</td>
<td>7.68</td>
<td>1.97</td>
<td>0.18</td>
<td>1.00</td>
<td>10.00</td>
</tr>
</tbody>
</table>

**Self-efficacy for writing a lesson plan**

Table 4.13 reports the adjusted descriptive data for self-efficacy for writing a
lesson plan after treatment (PostSE). The mean scores were adjusted using a formula
that included Cronbach’s Alpha to account for unreliability of using nonrandomized
subjects.

<p>| Table 4.13 | Adjusted Descriptive Data for Self-efficacy After Treatment (PostSE) |</p>
<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Adj. Mean</th>
<th>Adj. Std. Deviation</th>
<th>Adj. Std. Error</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>39</td>
<td>8.59</td>
<td>.84</td>
<td>.13</td>
<td>6.52</td>
<td>9.85</td>
</tr>
<tr>
<td>R</td>
<td>36</td>
<td>8.18</td>
<td>1.02</td>
<td>.17</td>
<td>5.76</td>
<td>9.73</td>
</tr>
<tr>
<td>RE</td>
<td>39</td>
<td>7.95</td>
<td>1.40</td>
<td>.22</td>
<td>2.85</td>
<td>9.85</td>
</tr>
<tr>
<td>Total</td>
<td>114</td>
<td>8.24</td>
<td>1.13</td>
<td>.11</td>
<td>2.85</td>
<td>9.85</td>
</tr>
</tbody>
</table>
ANCOVA

Table 4.14, reports the results of ANCOVA.

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Seq. Sum of Squares</th>
<th>Adj. Sum of Squares</th>
<th>Adj. Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>PreSE</td>
<td>1</td>
<td>57.80</td>
<td>59.47</td>
<td>59.47</td>
<td>*83.38</td>
</tr>
<tr>
<td>Treatment</td>
<td>2</td>
<td>9.70</td>
<td>9.70</td>
<td>4.85</td>
<td>*6.8</td>
</tr>
<tr>
<td>Error</td>
<td>110</td>
<td>78.46</td>
<td>78.46</td>
<td>.71</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>113</td>
<td>145.97</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p≤.005

To facilitate proper interpretation, the assumptions for ANCOVA were addressed. The covariate (PreSE) was administered before the treatment. The dependent variable, PostSE is normally distributed in the population for C and RE, but shows a slight negative skewness for R. The scores of different participants are statistically independent of one another. The population variance of PostSE is equal for all the groups. In addition, homogeneity of regression slopes is assumed.

The reliability-corrected ANCOVA result for PostSE covarying PreSE was significant for differences among treatment groups $F(2, 110) = 6.8, p=.002$. To find which pairs of treatments showed the greatest difference in means, the Tukey follow-up test was used to discover the difference between adjusted mean scores for control (C) and rubric with explanation (RE) was significant. C group had higher post-task self-efficacy scores ($M= 8.59$) than the RE group ($M= 7.95$). The results rejected the null hypothesis $H_01 (C=R=RE)$ at the .05 level, and led to the answer to research question 1, that there was significant difference in the means for self-efficacy post-treatment.
Post Treatment Achievement

Lesson plan 2

Table 4.15 reports the descriptive data for LP2. The mean scores were adjusted using a formula (see Chapter 3, Data Analysis) that included Cronbach’s Alpha to account for unreliability when using nonrandomized subjects.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Adj. Mean</th>
<th>Adj. Std. Deviation</th>
<th>Adj. Std. Error</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>39</td>
<td>30.38</td>
<td>6.64</td>
<td>1.06</td>
<td>14.29</td>
<td>46.21</td>
</tr>
<tr>
<td>R</td>
<td>36</td>
<td>35.53</td>
<td>8.73</td>
<td>1.46</td>
<td>7.24</td>
<td>45.79</td>
</tr>
<tr>
<td>RE</td>
<td>39</td>
<td>32.22</td>
<td>8.68</td>
<td>1.39</td>
<td>12.22</td>
<td>47.87</td>
</tr>
<tr>
<td>Total</td>
<td>114</td>
<td>32.67</td>
<td>8.26</td>
<td>.933</td>
<td>7.24</td>
<td>47.87</td>
</tr>
</tbody>
</table>

ANCOVA

ANCOVA was employed to partial out the effects of LP1 on lesson plan 2 (LP2). To facilitate proper interpretation, the assumptions for ANCOVA were addressed. The covariate (LP1) was administered before the treatment. The dependent variable, LP2 is normally distributed in the population for C and R, but is negatively skewed for RE. The scores of different participants are statistically independent of one another. The population variance of LP2 is equal for all the groups. In addition, homogeneity of regression slopes is assumed.
Table 4.16, reports the results of ANCOVA for LP2.

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Seq. Sum of Squares</th>
<th>Adj. Sum of Squares</th>
<th>Adj. Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP1</td>
<td>1</td>
<td>807.16</td>
<td>524.32</td>
<td>524.32</td>
<td>8.63</td>
<td>.004</td>
</tr>
<tr>
<td>Treatment</td>
<td>2</td>
<td>221.49</td>
<td>221.49</td>
<td>110.75</td>
<td>1.82</td>
<td>.166</td>
</tr>
<tr>
<td>Error</td>
<td>110</td>
<td>6683.16</td>
<td>6683.16</td>
<td>60.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>113</td>
<td>7711.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The reliability-corrected ANCOVA was not significant for treatment ($p=.166$).

The result retains the null hypothesis $H_{02}$ and led to the answer to research question 2, that there were no significant differences among treatment groups in achievement on LP2 while accounting for the influence of LP1.

**Survey Results: Post Treatment Items**

The results of the data analysis of the survey items that may have been affected by the treatment (4, 8, and 9a-f) are presented below. The chi-square test of independence was used to test for differences among groups for item 4, participants’ time on task, in order to answer supplemental research question 1, and for item 8, overall perception of the usefulness of the rubric in order to answer supplemental research question 2. For item 9a-f, chi-square was employed to test for group differences in order to answer supplemental research questions 3a–3f.

**Time on task**

To examine possible differences in time spent on the lesson plan task, survey item 4 asked participants for an estimated “…total amount of time you estimate you spent working on creating the lesson plan.” After examining the frequency data, the
five response choices were collapsed into three categories for chi-square analysis. The two shortest time choices were made “1 – 3 hours,” while “3 – 5 hours” was not collapsed, and “5 or more hours” included the infrequent response “more than 7 hours.” Table 4.17 shows the results.

Table 4.17  Chi Square Analysis of Time Spent Creating Lesson Plan by Treatment

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Item 4. Estimated Amount of Time Spent Creating Lesson Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1–3 hours</td>
</tr>
<tr>
<td>Control (C)</td>
<td></td>
</tr>
<tr>
<td>(n=39)</td>
<td>Count</td>
</tr>
<tr>
<td></td>
<td>% of group</td>
</tr>
<tr>
<td></td>
<td>% of item</td>
</tr>
<tr>
<td>Rubric (R)</td>
<td></td>
</tr>
<tr>
<td>(n=36)</td>
<td>Count</td>
</tr>
<tr>
<td></td>
<td>% of group</td>
</tr>
<tr>
<td></td>
<td>% of item</td>
</tr>
<tr>
<td>Rubric+ (RE)</td>
<td></td>
</tr>
<tr>
<td>(n=39)</td>
<td>Count</td>
</tr>
<tr>
<td></td>
<td>% of group</td>
</tr>
<tr>
<td></td>
<td>% of item</td>
</tr>
<tr>
<td>Total</td>
<td>Count</td>
</tr>
<tr>
<td>(n=114)</td>
<td>% of total</td>
</tr>
</tbody>
</table>

Chi-Square = 0.675, DF = 4, P-Value = 0.954
3 cells with expected counts less than 5.0

The majority of all participants (68.75%) reported spending 1–3 hours creating the lesson plan and this pattern of response is consistent for each treatment group. The data among groups (% of item) are consistent for “1-3 hours” and “3-5 hours”. Likewise, the count for “5+ hours” reveals comparatively small numbers among groups.
result of the chi-square test of independence confirmed there was not a significant relationship among time spent creating lesson plans and treatment ($\chi^2 = 0.675, df = 4, p = 0.954$) and lead to the answer for research question 1, that neither rubric use nor explanation of the rubric lead to differences in the amount of time spent creating lesson plan 2 (LP2).

**Perceptions of Using the Rubric**

The following reports the data analysis results for survey items 8 and 9a-f, which focused on the participants’ perceptions of using the rubric. Responses included only two treatment groups, R and RE. The control group was excluded because it did not receive a rubric. The results are organized by survey item sequence.

**Overall usefulness of the rubric**

Table 4.18 shows the chi-square analysis for survey item 8, “Indicate how useful you found the rubric overall while you were developing your lesson plan.” After examining the frequency data, the five response choices were collapsed into four categories for chi-square analysis, “not at all useful,” “a little useful,” “somewhat useful,” and “very useful.” The last category included the few responses for “extremely useful.”
Table 4.18  Chi Square Analysis of Overall Usefulness of Rubric by Treatment

<table>
<thead>
<tr>
<th>Item 8. Overall Usefulness of Rubric While Developing Lesson Plan</th>
<th>Not at all</th>
<th>Little</th>
<th>Somewhat</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rubric (R) Count (n=27)</td>
<td>3</td>
<td>13</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>% of total</td>
<td>11.11</td>
<td>48.15</td>
<td>29.63</td>
<td>11.11</td>
</tr>
<tr>
<td>% of item</td>
<td>42.86</td>
<td>54.17</td>
<td>34.78</td>
<td>60.00</td>
</tr>
<tr>
<td>Rubric+ (RE) Count (n=32)</td>
<td>4</td>
<td>11</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>% of total</td>
<td>12.50</td>
<td>34.37</td>
<td>46.88</td>
<td>6.25</td>
</tr>
<tr>
<td>% of item</td>
<td>57.14</td>
<td>45.83</td>
<td>65.22</td>
<td>40.00</td>
</tr>
<tr>
<td>Total Count (n=59)</td>
<td>7</td>
<td>24</td>
<td>23</td>
<td>5</td>
</tr>
<tr>
<td>% of total</td>
<td>11.86</td>
<td>40.68</td>
<td>38.98</td>
<td>8.47</td>
</tr>
</tbody>
</table>

Chi-Square = 2.232, DF = 3, P-Value = 0.526
4 cells with expected counts less than 5.0

With respect to the usefulness of the rubric within groups (% of group), the combined percentage of “little” and “somewhat” for R (77.78%) and RE (81.25%) were nearly the same. The result of the chi-square test of independence confirmed there was not a significant relationship between usefulness of the rubric and treatment ($\chi^2 = 2.232$, $df = 3$, $p = 0.526$). The two levels of explanation about rubric use did not lead to significant differences in participants’ perceptions of the usefulness of the rubric.

**Helpfulness of rubric for explaining**

Survey item 9 has six parts (lettered a–f) and used the Likert-type scale, shown in Figure 4.1 below. These items focused on the participants’ perceptions of using the
rubric and instructed them to indicate their level of agreement to different statements, such as whether the rubric provided the participant with explanations, helped understanding, helped establish goals, provided feedback, saved time, and helped with self-evaluation. The Likert-type scale data were analyzed as ordinal data. Items 9a, 9c, and 9f were written with a positive implication and to prevent patterned responses, items 9b, 9d, and 9e were written with a negative implication.

<table>
<thead>
<tr>
<th>A = strongly agree</th>
<th>B = agree</th>
<th>C = undecided</th>
<th>D = disagree</th>
<th>E = strongly disagree</th>
</tr>
</thead>
</table>

Table 4.19 below, shows chi square analysis for survey item 9a, in which participants responded with their level of agreement to the statement, “The rubric I was given…Helped me by providing explanations for how to do the lesson plan before I began working on it.” No participant selected “strongly disagree.” The overall high response of “agree” (37.33%) is similar to the two treatment groups, R (30.56%) and RE (43.59%). Taken with the 18.67% overall response of “strongly agree,” the results indicate most participants agreed that the rubric helped by providing explanations for how to do the lesson plan. The chi square test of independence could not be run because the condition of expected cell counts were not met.
Table 4.19  Chi Square Analysis of Item 9a
Rubric Provided Explanation for How to Do the Lesson Plan

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Count</th>
<th>% of group</th>
<th>% of item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rubric (R)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=36)</td>
<td>8</td>
<td>22.22</td>
<td>50.00</td>
</tr>
<tr>
<td>% of group</td>
<td>8</td>
<td>22.22</td>
<td>57.14</td>
</tr>
<tr>
<td>% of item</td>
<td>11</td>
<td>30.56</td>
<td>39.29</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>25.00</td>
<td>60.00</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Rubric+ (RE)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=39)</td>
<td>8</td>
<td>20.51</td>
<td>50.00</td>
</tr>
<tr>
<td>% of group</td>
<td>6</td>
<td>15.38</td>
<td>42.86</td>
</tr>
<tr>
<td>% of item</td>
<td>17</td>
<td>43.59</td>
<td>60.71</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>15.38</td>
<td>40.00</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5.13</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=75)</td>
<td>16</td>
<td>21.33</td>
<td>37.33</td>
</tr>
<tr>
<td>% of total</td>
<td>14</td>
<td>18.67</td>
<td>20.00</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>37.33</td>
<td>2.67</td>
</tr>
</tbody>
</table>

Chi-Square = 4.058, DF = 4, P-Value = 0.526
1 cells with expected counts less than 1.0, 2 cells with expected counts less than 5.0

Helpfulness of rubric for understanding the requirements

Table 4.20 below, shows chi square analysis for survey item 9b, in which participants responded with their level of agreement to the statement, “The rubric I was given…Did not help me understand the requirements to earn the highest possible score on the lesson plan assignment.” This survey item was written with a negative implication to prevent patterned responses.

The overall high response was “disagree” (33.33%) and “strongly disagree” (24%). The within groups pattern is similar. The results indicate the majority of
participants perceived the rubric helped them to understand the requirements on the lesson plan assignment. The chi square test of independence could not be run because the condition of expected cell counts were not met.

<table>
<thead>
<tr>
<th>Table 4.20</th>
<th>Chi Square Analysis of Item 9b</th>
<th>Rubric Did Not Help Understanding of Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>missing</td>
<td>A strongly agree</td>
</tr>
<tr>
<td>Treatment</td>
<td>Count</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rubric (R)</td>
<td>Count</td>
<td>8</td>
</tr>
<tr>
<td>(n=36)</td>
<td>% of group</td>
<td>22.22</td>
</tr>
<tr>
<td></td>
<td>% of item</td>
<td>50.00</td>
</tr>
<tr>
<td>Rubric+ (RE)</td>
<td>Count</td>
<td>8</td>
</tr>
<tr>
<td>(n=39)</td>
<td>% of group</td>
<td>20.51</td>
</tr>
<tr>
<td></td>
<td>% of item</td>
<td>50.00</td>
</tr>
<tr>
<td>Total</td>
<td>Count</td>
<td>16</td>
</tr>
<tr>
<td>(n=75)</td>
<td>% of total</td>
<td>21.33</td>
</tr>
</tbody>
</table>

Chi-Square = 2.108, DF = 5
2 cells with expected counts less than 1.0, 6 cells with expected counts less than 5.0
Helpfulness of rubric for establishing goals

Table 4.21 shows chi square analysis for survey item 9c, in which participants responded with their level of agreement to the statement, “The rubric I was given...Helped me establish the goals for creating my lesson plan.” The overall data shows that 40% of the participants responded with “agree” and 20.51% responded with “strongly agree.” The results indicate the majority of participants responded with a high level of agreement that the rubric helped them establish goals for creating the lesson plan. The chi square test of independence could not be run because the condition of expected cell counts were not met.

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Rubric (R)</th>
<th>Count</th>
<th>missing</th>
<th>A strongly agree</th>
<th>B agree</th>
<th>C undecided</th>
<th>D disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n=36)</td>
<td></td>
<td>8</td>
<td>9</td>
<td>12</td>
<td>7</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>% of group</td>
<td>22.22</td>
<td>25.00</td>
<td>33.33</td>
<td>19.44</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>% of item</td>
<td>50.00</td>
<td>52.94</td>
<td>40.00</td>
<td>63.64</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Rubric+ (RE)</td>
<td></td>
<td>8</td>
<td>8</td>
<td>18</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(n=39)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of group</td>
<td>20.51</td>
<td>20.51</td>
<td>46.15</td>
<td>10.26</td>
<td>2.56</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>% of item</td>
<td>50.00</td>
<td>47.06</td>
<td>60.00</td>
<td>36.36</td>
<td>100.00</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>16</td>
<td>17</td>
<td>30</td>
<td>11</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(n=75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of total</td>
<td>21.33</td>
<td>22.67</td>
<td>40.00</td>
<td>14.67</td>
<td>1.33</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

Chi-Square = 2.962, DF = 4
2 cells with expected counts less than 1.0, 2 cells with expected counts less than 5.0
Rubric provides feedback

Table 4.22 shows chi square analysis for survey item 9d, in which participants responded with their level of agreement to the statement, “The rubric I was given…Did not provide me understandable feedback on how well I was completing the lesson plan.” This survey item was written with a negative implication to prevent patterned responses. The chi square test of independence could not be run because the condition of expected cell counts were not met. The overall response favors agreement.

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Rubric (R)</th>
<th>Count (n=36)</th>
<th>Rubric+ (RE)</th>
<th>Count (n=39)</th>
<th>Total (n=75)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>A strongly agree</td>
<td>B agree</td>
<td>C undecided</td>
<td>D disagree</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>8</td>
<td>0</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>% of group</td>
<td>22.22</td>
<td>—</td>
<td>22.22</td>
<td>13.89</td>
</tr>
<tr>
<td></td>
<td>% of item</td>
<td>50.00</td>
<td>—</td>
<td>72.73</td>
<td>38.46</td>
</tr>
<tr>
<td></td>
<td>Count</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>% of group</td>
<td>20.51</td>
<td>2.56</td>
<td>7.69</td>
<td>20.51</td>
</tr>
<tr>
<td></td>
<td>% of item</td>
<td>50.00</td>
<td>100.00</td>
<td>27.27</td>
<td>61.54</td>
</tr>
<tr>
<td></td>
<td>Count</td>
<td>16</td>
<td>1</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>% of total</td>
<td>21.33</td>
<td>1.33</td>
<td>14.67</td>
<td>17.33</td>
</tr>
</tbody>
</table>

Chi-Square = 4.358, DF = 5
2 cells with expected counts less than 1.0, 6 cells with expected counts less than 5.0
Rubric saves time

Table 4.23 below, shows chi square analysis for survey item 9e, in which participants responded with their level of agreement to the statement, “The rubric I was given…Did not save me time in completing the assignment.” This survey item was written with a negative implication. Overall response is balanced between “undecided” (26.67%) and “disagree” (24%). The chi square test of independence confirms there was not a significant relationship between groups ($\chi^2 = 0.757, df = 5, p = 0.980$).

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Count</th>
<th>% of group</th>
<th>% of item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rubric (R)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=36)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Count</td>
<td>8</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>% of group</td>
<td>22.22</td>
<td>5.56</td>
<td>11.11</td>
</tr>
<tr>
<td>% of item</td>
<td>50.00</td>
<td>40.00</td>
<td>44.44</td>
</tr>
<tr>
<td>Rubric+ (RE)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=39)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Count</td>
<td>8</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>% of group</td>
<td>20.51</td>
<td>7.69</td>
<td>12.82</td>
</tr>
<tr>
<td>% of item</td>
<td>50.00</td>
<td>60.00</td>
<td>55.56</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=75)</td>
<td>16</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>% of total</td>
<td>21.33</td>
<td>6.67</td>
<td>12.00</td>
</tr>
</tbody>
</table>

Chi-Square = 0.757, DF = 5, P-Value = 0.98
6 cells with expected counts less than 5.0
Rubric helps self-evaluation

Table 4.24 below, chi square analysis for survey item 9f, in which participants responded with their level of agreement to the statement, “The rubric I was given...Helped me by providing information I could use to self-evaluate my work in creating the lesson plan.” No participant selected response items, “strongly disagree,” or “disagree,” so these two items do not appear in the table. Overall response was “agree” (40%) and “strongly agree” (22.67%) with R and RE following a similar pattern. The results indicate participants perceived the rubric helped provide information for self-evaluation. The chi square test of independence confirms there was not a significant relationship between groups ($\chi^2 = 0.406$, $df = 3$, $p = 0.939$).

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Count</th>
<th>% of group</th>
<th>% of item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rubric (R) (n=36)</td>
<td>8</td>
<td>22.22</td>
<td>50.00</td>
</tr>
<tr>
<td>Rubric+ (RE) (n=39)</td>
<td>8</td>
<td>20.51</td>
<td>50.00</td>
</tr>
<tr>
<td>Total (n=75)</td>
<td>16</td>
<td>21.33</td>
<td>40.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>missing</th>
<th>A strongly agree</th>
<th>B agree</th>
<th>C undecided</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>9</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td>% of group</td>
<td>25.00</td>
<td>38.89</td>
<td>13.89</td>
</tr>
<tr>
<td>% of item</td>
<td>52.94</td>
<td>46.67</td>
<td>41.67</td>
</tr>
</tbody>
</table>

Chi-Square = 0.406, DF = 3, P-Value = 0.939
Correlations Between Measures

Table 4.25 contains the Pearson product-moment correlation coefficients calculated to describe the relationships between the participants’ individual pre- and post self-efficacy score (PreSE and PostSE) and the mean scores on lesson plan 1 (LP1) and lesson plan 2 (LP2). The findings revealed a positive moderate relationship ($r = .63$), significant at the .000 level, between the pre-treatment self-efficacy (PreSE) and post-treatment self-efficacy scores (PostSE). No significant relationship was found between PostSE mean score and achievement on either LP1 or LP2. The findings also revealed no significant relationship between participants’ scores on LP1 and LP2.

<table>
<thead>
<tr>
<th></th>
<th>PreSE</th>
<th>PostSE</th>
<th>LP1</th>
</tr>
</thead>
<tbody>
<tr>
<td>PostSE</td>
<td>0.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LP1</td>
<td>0.18</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>LP2</td>
<td>0.03</td>
<td>0.08</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table 4.26 below, shows correlation data derived from the survey regarding participants’ perceptions of using the rubric in this study.

<table>
<thead>
<tr>
<th></th>
<th>Item 9a</th>
<th>Item 9c</th>
<th>Item 9d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RUBEXPLAIN</td>
<td>RUBGOAL</td>
<td>RUBFEED</td>
</tr>
<tr>
<td>Item 9c</td>
<td>RUBGOAL</td>
<td>0.626**</td>
<td></td>
</tr>
<tr>
<td>Item 9d</td>
<td>RUBFEED</td>
<td>0.497**</td>
<td>0.272</td>
</tr>
<tr>
<td>Item 9f</td>
<td>RUBEVAL</td>
<td>0.539**</td>
<td>0.653**</td>
</tr>
</tbody>
</table>

** $p \leq .000$
The findings revealed a positive moderate relationship \( (r = .626) \) between RUBGOAL and the perception the rubric was helpful in providing explanations of how to do the lesson plan before the task was begun (RUBEXPLAIN). The data shows a positive low correlation \( (r = .497) \) between RUBEXPLAIN and the rubric provided feedback on how well the participant was completing the lesson plan (RUBFEED). A positive moderate relationship also exists between the perception the rubric provided information the participant could use to self-evaluate their lesson plan (RUBEVAL) and RUBEXPLAIN \( (r = .539) \) and RUBGOAL \( (r = .653) \).

**Summary of Hypotheses Tested**

The following is a brief report on the findings related to each of the null hypotheses tested in this study.

**Hypothesis 1.** There will be no significant difference in the means for post-treatment self-efficacy among three treatment groups.

The above null hypothesis was rejected at the .05 level, because the control group (C) exhibited higher self-efficacy after the treatment. More specifically, the R and RE groups did not gain significantly in self-efficacy despite receiving treatments.

**Hypothesis 2.** There will be no significant difference in the means for achievement for writing a lesson plan among three treatment groups, no rubric, rubric given without explanation, and rubric given with explanation.

The above null hypothesis was retained at the .05 level.
Summary of Supplemental Research Questions

Table 4.27 summarizes the findings for supplemental research questions 1 and 2.

Table 4.28 summarizes the findings of supplemental research questions 3a–3f.

Table 4.27  Summary of Findings for Supplemental Research Questions 1 and 2

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1. Does the strategy for implementation of rubric use lead to differences in the amount of time spent creating the second lesson plan?</td>
<td>Chi square; no significant differences</td>
</tr>
<tr>
<td>Q2. Do the two strategies for implementation of rubric use lead to differences in participants’ perceptions of the usefulness of the rubric?</td>
<td>Chi square; no significant differences</td>
</tr>
</tbody>
</table>
Table 4.28  Summary of Findings for Supplemental Research Questions 3a-3f

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q3a. Does the strategy for implementation of rubric use lead to differences in participants' perceptions of how to write a lesson plan before starting the task?</td>
<td>Chi square conditions not met</td>
</tr>
<tr>
<td>Q3b. Does the strategy for implementation of rubric use lead to differences in participants' understanding of the requirements to earn the highest possible score on the lesson plan assignment?</td>
<td>Chi square conditions not met</td>
</tr>
<tr>
<td>Q3c. Does the strategy for implementation of rubric use lead to differences in participants' perceptions of the helpfulness of the rubric for setting goals to create the lesson plan?</td>
<td>Chi square conditions not met</td>
</tr>
<tr>
<td>Q3d. Does the strategy for implementation of rubric use lead to differences in participants' perceptions of understandable feedback derived from using the rubric?</td>
<td>Chi square conditions not met</td>
</tr>
<tr>
<td>Q3e. Does the strategy for implementation of rubric use lead to different amounts of time saved completing the lesson plan writing assignment?</td>
<td>Chi square; no significant differences</td>
</tr>
<tr>
<td>Q3f. Does the strategy for implementation of rubric use lead to differences in participants' perceptions of using the rubric for self-evaluation while creating the lesson plan?</td>
<td>Chi square; no significant differences</td>
</tr>
</tbody>
</table>
CHAPTER 5
DISCUSSION

Overview of the Study

The primary purpose of this study was to investigate the effect of three strategies for implementation of an analytic trait, task-specific rubric, no rubric (C), rubric given without explanation (R), and rubric given with explanation (RE), on teacher credential candidates’ self-efficacy and achievement for writing a lesson plan. Specifically, the research questions were:

1. Do the three strategies for analytic trait, task-specific rubric use lead to differences in the means for posttask self-efficacy for writing a lesson plan?
2. Do the three strategies for analytic trait, task-specific rubric use lead to differences in the quality of lesson plans written by subjects in the three treatment groups?

The study was implemented in the Fall of 2005 and Spring of 2006 at a large, west coast, public university with 114 volunteer participants recruited from introductory courses for teacher credential candidates. Table 5.1 shows the null hypotheses.

<table>
<thead>
<tr>
<th>Table 5.1</th>
<th>Hypotheses Tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{01}$</td>
<td>There will be no significant difference in the means for post-task self-efficacy for writing a lesson plan among three treatment groups, no rubric, rubric given without explanation, and rubric given with explanation.</td>
</tr>
<tr>
<td>$H_{02}$</td>
<td>There will be no significant difference in the means for achievement for writing a lesson plan among three treatment groups, no rubric, rubric given without explanation, and rubric given with explanation.</td>
</tr>
</tbody>
</table>
The hypothesis testing resulted in rejecting $H_01$ and retaining $H_02$.

The following interpretation and discussion of the data analysis results are presented for each hypothesis and then the supplemental research questions. Finally, the implications for educators and recommendations for future research are addressed.

**Interpretation of Hypothesis Testing**

**Hypothesis 1**

$H_01$ There will be no significant difference in the means for post-task self-efficacy for writing a lesson plan among three treatment groups, no rubric, rubric given without explanation, and rubric given with explanation.

This hypothesis was rejected. The results of the reliability-corrected ANCOVA ($F = 6.8, p = .002$) for the post-task self-efficacy measure (PostSE) revealed that the control group, C, had a significantly higher self-efficacy score after treatment ($M = 8.59$) than both the rubric given without explanation group, R ($M=8.18$) and the rubric given with explanation group, RE ($M=7.95$). The overall mean post-task self-efficacy (PostSE) score was 8.24 on a 10-point scale indicating a high degree of self-efficacy for all groups. The complete ranges of mean scores for self-efficacy were categorized as low level if the mean was in the range of 1–3, as middle level for means within 4-6, and as high level for means within 7-10. The results regarding self-efficacy for R, rubric without explanation group, and RE, rubric with explanation, fits with the mixed results found in other research for the effect of a rubric on boosting confidence or self-efficacy. Research suggests increases to self-efficacy can be facilitated via a rubric, namely by providing performance standards (Schunk, 2001), by rubric-referenced self-assessment (Paris &
Paris, 2001) and self-evaluation (Schunk & Ertmer, 1999). In one study by Andrade and Du (2005), participants’ self-report data indicated an increased confidence when using a rubric. By contrast, in a more recent study, Andrade (2006a) did not find particular evidence to support the claim rubrics provide an increase in learners’ self-efficacy.

Given a high mean self-efficacy for each group, the data indicates the participants felt confident about writing a lesson plan. Participants may have already known or believed they knew how to write a lesson plan and the introduction of a rubric may have interfered with increasing self-efficacy for R and RE. If interference by the rubric accounts for any portion of the result for R and RE, it may be attributed to two factors reported by course instructors. All course instructors did not adopt the model of lesson planning the rubric was designed to evaluate and grades given by the instructors were not related to their score on the rubric. This explanation implies participants in R and RE may have experienced inconsistency between using the rubrics and the grading standard. In addition, participants would have had little or no motivation to use the rubric. The inconsistency and low motivation would have been counterproductive for increasing self-efficacy, which is related to a person knowing what to do and perceiving the challenge worth accomplishing (Pintrich & Schunk, 2002).

The methodology of the study is likely a strong influence on the result of PostSE. The nonrandom assignment of participants and matching of intact classes to treatments may have allowed for groups to have significant numbers of participants with skills, abilities, or attitudes that for C group, enabled their self-efficacy and for R and RE, hindered it. In the naturalistic setting of this study, it is likely that factors that are known to increase self-efficacy may have been introduced outside the context of the rubric. For example, the course instructors of the control groups may have given
personal feedback about participants’ particular performance that would be considered a more influential means of feedback and therefore more likely to increase self-efficacy (Zimmerman, Bandura, & Martinez-Pons, 1992) than the generic, descriptive phrases a rubric offers.

Hypothesis 2

$H_0^2$ There will be no significant difference in the means for achievement for writing a lesson plan among three treatment groups, no rubric, rubric given without explanation, and rubric given with explanation.

This hypothesis was retained. The reliability-corrected ANCOVA revealed there were no significant differences among groups in post treatment achievement in writing a second lesson plan (LP2), while accounting for the influence of LP1. This result indicates the rubric may not have stimulated a discernable change in participants’ achievement for writing a lesson plan that could not be attributed to the influences outside the control of this study, maturity, or even when the influence of experience from LP1 was statistically controlled.

These findings are not consistent with a recent study by Andrade (2006b) in which elementary students’ scores on written stories or essays were higher when they used a rubric to self-assess their rough drafts. Another study (Andrade, 1999) found the effects of instructional rubrics on 8th graders’ writing skills resulted in significant differences for writing an essay.

Five possible explanations for insignificant results for null hypothesis 2 are worth noting. First, the manner in which the treatments were administered may not have provided enough incentive for participants to take the study seriously and use the rubric to self-assess their progress in writing lesson plans. Despite the agreed upon
methods and protocols for handling their part of the study, course instructors for both
rubric without explanation group (R) and rubric with explanation group (RE) retained
control over how the rubrics were presented and more importantly, whether
participants were informed that the rubric would be used to evaluate their lesson plans
and subsequently, impact their grade on the lesson plan assignment. In the post-
treatment survey, items 8 and 9a-f were designed to collect data regarding participants’
perceptions of the usefulness of the rubric. In summary, the survey findings show the
majority of participants reported levels of agreement with the perception that the rubric
provided explanations for how to do the lesson plan, helped understanding of the
requirements, helped establish goals, and helped with self-evaluation. The participants’
perception of the rubric as reported on the survey however, does not indicate how the
perception was formed; whether it was formed from actual use or that the perception
was formed based on knowledge or speculation of its use.

The second explanation is that the length of time participants in RE were
exposed to the treatment may have been too short to have a noticeable effect. A single
20-30 minute explanation of the use of a rubric may not have been sufficient to facilitate
transfer of the new knowledge to actual performance on the assignment. Consistent
with recommendations from Andrade (2000, 1999), a longer, more involved
instructional intervention may have affected a change in participants’ strategies for
writing lesson plans. While simply handing out and explaining a rubric seemed to help
improve writing for 8th graders (Andrade, 2001), it appeared in this study that a more
intensive instructional intervention may be necessary in order to help improve
performance for teacher credential candidates.

Alternately, participants may have felt they knew enough about writing a lesson
plan or had sufficient capabilities to figure it out that any explanation on the use of a
rubric to help them accomplish the task was perceived as superfluous. The study did not capture enough data on participants’ actual use of the rubric to infer any relationship among use, explanation of use, and achievement.

This leads to the third possible reason for the results. The presentation and explanation on the use of the rubric may not have been adequately effective as instruction to produce a significant, positive change in achievement score for RE. The brief presentation was an explanation of the use of the rubric, but did not offer any generative activities for participants to have a meaningful learning experience with the content. Generative activities are described as learning processes that require the learner to actively make relationships between components of the new content being learned or between the new content and their existing knowledge. Examples of generative activities include, writing summaries, creating concept maps, or creating examples (Grabowski, 1996; Wittrock, 1990). Meaningful learning may have been more likely to occur in this study if RE participants had been given activities to apply and/or practice their new knowledge of rubrics to writing a lesson plan.

Fourth, the literature claims that rubrics provide feedback and thus inspire higher performance (Arter & McTighe, 2001), if the learner knows how to use the rubric in this manner (Andrade & Du, 2005). It may be that no significant differences in achievement were found, in part, because the participant perceived the rubric as merely guidance in how to craft the lesson plan, rather than as feedback on how they performed.

Last, the methodology must be considered as a possible factor in the result, as discussed previously. It was not clear from discussions with the RE course instructors the extent the rubric was promoted or not, as a tool to help with creating the lesson plan. The descriptions of the course instructors’ approaches to teaching lesson plans
reveals varying degrees of importance on the written plan. For example, Instructor A indicated that the written lesson was not as critical to course performance as the actual implementation. By contrast, Instructor C used a lesson planning model to instruct participants, taught five other models, and expected participants to use one particular method in developing their own lesson plan.

**Interpretation of Supplemental Research Questions**

Additional descriptive research questions were posed to seek the degree of relationship that exists between the independent variable, strategies for implementation of rubrics, and participants’ self-report on time spent writing a lesson plan and perceptions of the usefulness of the rubric. The data were collected via the survey administered immediately after the study. The research questions are restated below and followed by discussion. A summary of all participant perceptions is at the end of this section.

**Q1: Time on Task**

1. Does the strategy for implementation of analytic trait, task-specific rubric use, no rubric, rubric given without explanation, or rubric given with explanation, lead to differences in the amount of time spent creating the second lesson plan?

The chi-square test confirmed no significant differences. The descriptive data revealed a nearly equal percentage among groups for responses to each choice. The overwhelming high percentage for each group was 1-3 hours. This is not consistent with findings by Andrade (2005), where students using rubrics saved time on task. It may be
that the rubric may not have helped R and RE to spend less time on task than the control.

Q2: Perception of Rubric Usefulness

2. Do the two strategies for implementation of analytic trait, task-specific rubric use, rubric given without explanation or rubric given with explanation, lead to differences in participants’ perception of the usefulness of the rubric?

The chi-square test revealed no significant differences in perception of the usefulness of the rubric between R, rubric without explanation, and RE, rubric with explanation. Given this and the research of Andrade and Du (2005) that found students who used rubrics believed there were positive outcomes, such as less anxiety about the assignment, improvements in the quality of their work, and fairer grades, more research appears warranted in this area.

Research question 3 has six parts (3a-f) corresponding to survey items 9a-f. The summary of findings for each part pertains only to R and RE, the groups receiving a rubric.

Q3a:

Strategy for Implementation of Rubric Helped to Explain How to Write a Lesson Plan

3a. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perception of how to write a lesson plan before starting the task?

Analysis of frequency data indicated that overall, participants’ perceptions favored agreement that the rubric helped them to know how to write the lesson plan. This finding is consistent with research in which students reported that rubrics helped
explain task objectives (Arter & McTighe, 2001). It may be that the criteria in the rubric provided participants with an outline of how to construct a lesson plan in terms of form and content. Associations between groups and response choices were not determined because the condition of expected cell counts for the chi square statistic were not met.

Q3b: Strategy for Implementation of Rubric Helped Understanding of Requirements

3b. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ understanding of the requirements to earn the highest possible score on the lesson plan assignment?

Analysis of frequency data indicated that overall, participants’ perceptions favored agreement that the rubric helped them understand the requirements to earn a high score. This finding lends some support to research in which students reported rubrics helped them understand task requirements (Whittaker, Salend, & Duhaney, 2001) and provided explanation of task objectives (Arter & McTighe, 2001). It may be that the description of the highest level of quality for each criterion listed in the rubric either explicitly helped participants’ or improved their perception of understanding how to earn the highest score possible. Associations between groups and response choices were not determined because the condition of expected cell counts for the chi square statistic were not met.

Q3c: Strategy for Implementation of Rubric Helped Set Goals

3c. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perception of the helpfulness of the rubric for setting goals to create the lesson plan?

Analysis of frequency data indicated that overall, participants’ perceptions favored a high level of agreement that the rubric helped them set goals for creating the
lesson plan. This finding is consistent with the literature on the benefit of a rubric to promote setting goals (Andrade, 2000; Arter & McTighe, 2001) and lends some support to the previous research of Andrade and Du (2005) that found students used rubrics as a guideline to plan an approach to an assignment. It may be that the criteria and levels of quality written in the rubric helped participants set goals for the lesson plan assignment. Associations between groups and response choices were not determined because the condition of expected cell counts for the chi square statistic were not met.

**Q3d: Strategy For Implementation of Rubric Provided Feedback**

3d. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of understandable feedback derived from using the rubric?

Analysis of frequency data showed a slightly favorable response toward agreement that the rubric provided understandable feedback. Associations between groups and response choices were not determined because the condition of expected cell counts for the chi square statistic were not met.

The finding of the participants’ perceptions lends modest support to the literature that claims rubrics provide students feedback on progress (Saddler & Andrade, 2002; Stevens & Levi, 2005). In a study conducted by Andrade and Du (2005), it was found that rubrics are not necessarily self-explanatory and students need to learn how to use them in order to take advantage of the benefits. It may be that R and RE had sufficient prior experience with lesson plans and/or rubrics that they did not need or perceive as much of a need to use the rubric for feedback. Considering the method of explanation of rubrics in this study, participants in RE may not have learned enough about how to effectively use the rubric for feedback to affect their perceptions.
Q3e: Strategy For Implementation of Rubric Saved Time

3e. Does the strategy for implementation of analytic trait, task-specific rubric use lead to different amounts of time saved completing the lesson plan writing assignment?

The chi square test of independence indicated there was not a significant relationship between groups. The frequency data indicates an overall balance between “undecided” and “disagree.” This suggests more research is needed to investigate the findings of Andrade (2005), who reported that students felt the rubric saved time by focusing their attention on what matters most for writing an essay. It may be that participants in both R and RE felt sufficient confidence in their ability to write the lesson plan without the use of a rubric. It may have been perceived that the time needed to use or learn the rubric would extend beyond their preferred amount of time for completing the task.

Q3f: Strategy for Implementation of Rubric Helps Self-Evaluation

3f. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of using the rubric for self-evaluation while creating the lesson plan?

The chi square test of independence indicated there was not a significant relationship between groups. The frequency data indicates a high level of agreement that the rubric helped self-evaluation. This finding is consistent with the literature on the benefit of a rubric to promote self-evaluation (Andrade & Du, 2005; Andrade, 2000; Arter & McTighe, 2001). In the Andrade and Du study it was reported students used the rubric for informal self-evaluation throughout the task development process. It seems
participants may have used the rubric to check their progress or assess the outcome of
the whole or a part of the lesson plan.

**Summary of Supplemental Research Question Findings**

Table 5.2 summarizes the interpretations of supplemental research questions 1
and 2. Table 5.3 summarizes the interpretations of supplemental research question 3a-
3f, regarding the participants’ perceptions of using the rubric.

<p>| Table 5.2  Summary of Supplemental Research Questions 1 and 2 with Interpretations |</p>
<table>
<thead>
<tr>
<th>Research Question</th>
<th>Finding</th>
<th>Brief Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in the amount of time spent creating the second lesson plan?</td>
<td>Chi square; no significant differences</td>
<td>Rubric use did not significantly decrease time on task for R and RE. More research needed to find consistency with Andrade (2005).</td>
</tr>
<tr>
<td>Q2. Do the two strategies for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of the usefulness of the rubric?</td>
<td>Chi square; no significant differences</td>
<td>Rubric use did not significantly affect perceptions of usefulness of the rubric. More research needed to find consistency with Andrade and Du (2005).</td>
</tr>
</tbody>
</table>

Findings for research questions 1 and 2 are not consistent with previous research cited in Table 5.2. More research is warranted to investigate previous findings and substantiate claims that rubrics can save time and enhance learners’ perceptions of the overall usefulness of a rubric for complex tasks.
<table>
<thead>
<tr>
<th>Research Question</th>
<th>Finding</th>
<th>Brief Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q3a. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of how to write a lesson plan before starting the task?</td>
<td>Chi square conditions not met</td>
<td>Participants’ responses favored agreement. This is consistent with Arter and McTighe (2001).</td>
</tr>
<tr>
<td>Q3b. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ understanding of the requirements to earn the highest possible score on the lesson plan assignment?</td>
<td>Chi square conditions not met</td>
<td>Participants’ responses favored agreement. This is consistent with Arter and McTighe (2001).</td>
</tr>
<tr>
<td>Q3c. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of the helpfulness of the rubric for setting goals to create the lesson plan?</td>
<td>Chi square conditions not met</td>
<td>Participants’ responses favored a high level of agreement. This is consistent with Andrade and Du (2005).</td>
</tr>
<tr>
<td>Q3d. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of understandable feedback derived from using the rubric?</td>
<td>Chi square conditions not met</td>
<td>Participants’ responses slightly favor agreement. More research needed to investigate findings of Saddler and Andrade (2002).</td>
</tr>
<tr>
<td>Q3e. Does the strategy for implementation of analytic trait, task-specific rubric use lead to different amounts of time saved completing the lesson plan writing assignment?</td>
<td>Chi square; no significant differences</td>
<td>Participants’ responses inconclusive. More research is needed to support findings of Andrade (2005).</td>
</tr>
<tr>
<td>Q3f. Does the strategy for implementation of analytic trait, task-specific rubric use lead to differences in participants’ perceptions of using the rubric for self-evaluation while creating the lesson plan?</td>
<td>Chi square; no significant differences</td>
<td>Participants’ responses highly favor agreement. This is consistent with Andrade and Du (2005).</td>
</tr>
</tbody>
</table>
Findings for research questions 3a, 3b, 3c, 3d, 3f indicated slightly differing levels of participant responses to “agree” with the research questions. These findings regarding participants’ perceptions of the rubric offer modest support for previous research, as cited. Question 3e was inconclusive. More research is needed to investigate learners’ perceptions of using a rubric for different purposes on different complex tasks in various educational contexts.

This study found participants had high self-efficacy for the task of writing a lesson plan, which indicates a high degree of confidence in their ability to accomplish the task. High self-efficacy may have been a strong influence on participants’ perceptions of the usefulness of the rubric. It may also be that participants in this study were not good judges of their own perceptions or abilities. Another possible explanation is that participants who are post-baccalaureate, teacher credential candidates, are not affected by the use of a rubric. At this level of education, participants may feel confident in their abilities to perform academic tasks and therefore, may be less likely to seek external support for their work, such as that offered by a rubric. In support of this position, this study found participants reported limited use of rubrics for evaluating their own work or for evaluating their students’ work. Given limited use of rubrics prior to the study, participants may not have been inclined to fully use the rubric in the context of the study.

Limitations of the Study

The reading of this study should be done with an understanding of the limitations with respect to recruitment of participants, assignment of treatments, procedures, raters, instrumentation, characteristics of participants, and nature of task. This study was conducted in a naturalistic educational context in which all the
participants and their course instructors interacted with one another and the study materials in a manner that was as much a part of the established course curriculum as possible. Because this may have introduced extraneous variables outside the control of the study, generalization of these research findings should be done with caution. A detailed description of the limitations is included in Chapter 3.

**Implications for Educators**

Although this research study did not offer significant support for the claims that rubrics can promote self-efficacy or student achievement, more research is needed. Other research studies indicate that rubrics continue to hold promise as instruments to promote academic achievement when accompanied with instruction and guidance in their use with self-assessment (Andrade & Du, 2005; Stevens & Levi, 2005; Whittaker, Salend, & Duhaney, 2001). Rubrics may be a new instructional strategy for both teachers and students. Similar to teaching any new strategy, the knowledge and skills needed to effectively utilize the strategy must be developed over time with practice, application, and feedback.

A learners inclination to use a new strategy is partially dependent on his or her ability to self-regulate their learning experience. A key aspect of effective self-regulation is self-assessment, which has been shown effective for improving achievement (Pintrich & Schunk, 2002). A rubric can be used to teach students how to self-assess their work and therefore, should be effective for improving achievement (Andrade & Du, 2005; Andrade 2000, 2005). In addition, building confidence in students has been shown effective for improving achievement (Schunk, 1991; Zimmerman, Bandura, & Martinez-Pons, 1992) and using rubrics to build confidence in students’ perceptions of ability to effectively handle a comprehensive task should also prove effective.
In the context of this study, findings show the analytic trait, task-specific rubric may not have been as effective as the literature claims because the participants’ levels of confidence for writing the lesson plan and self-regulatory behavior may have been high. The results regarding self-efficacy, which were used to answer research question 1, showed significant difference for the control group. One cautionary implication for educators is to implement a rubric strategy in which there is high level of consistency between the rubric and the task, as well as direct impact of the rubric on the learners’ scores for the task.

One trait that has been shown to enhance self-efficacy for a task is the perception of knowing what to do and how to accomplish the task (Schunk & Ertmer, 2000). It may be that participants had such high levels of self-regulation, that the addition of a new strategy for accomplishing a task was not perceived as needed or necessary. Similarly, self-assessment may have been an established study trait for participants and that a different method was not perceived useful. Participants’ past experience with writing lesson plans may have been sufficient to provide enough confidence in their own ability to accomplish the task. The task used in this study may not have been perceived as complex enough that the participant felt they needed help with it.

Rubrics are reported most effective for challenging, complex, multi-dimensional tasks, in which the learner cannot easily see the path to a finished project (Arter & McTighe, 2001; Whittaker, Salend, et al. 2001). The value of a rubric is dependent on the learner’s perception of their self-efficacy, the ability to muster the resources necessary to accomplish a specific task. Results of the participants’ rating of their confidence in getting an “A” grade on the next lesson plan assignment showed a high level, labeled “pretty sure,” before and after treatment. It may be that if the task of writing a lesson plan were not perceived as sufficiently challenging, participants may not have been as
likely to seek additional assistance to accomplish it. Perceived ability to do a task is a key aspect of self-efficacy, so, it may also be that if participants believed they could accomplish the assignment the need for the rubric would have likely diminished. More research is needed to investigate the associations among rubrics, achievement on tasks, and problem solving methods for complex tasks.

The effectiveness of an instructional rubric strategy may depend on the characteristics of the student and the type of learning task. In this study, findings indicated participants had high self-efficacy to start and therefore may have been too mature in their academic ability for the rubric strategy to be effective. Also, the learning task of creating a lesson plan may not have been challenging enough or may have been perceived as well within the zone of proximal development of the participants to accomplish using their own strategies and resources.

**Recommendations for Future Research**

It was the intent of this study to investigate the use of an analytic trait, task-specific rubric as an instructional instrument to increase self-efficacy for the task of writing a lesson plan and to scaffold participants in the achievement of a complex task. The results do not lend unqualified support for the use of rubrics. Numerous methodological issues allowed for uncontrolled variables to influence the outcomes of this study, which may have caused it to be inconsistent with previous research. Lessons learned from this study should help future research continue to explore the conditions under which rubrics enhance student achievement and in particular, further explore the relationship between rubrics and self-efficacy. This general recommendation is partially based on personal correspondence with Dr. Heidi Andrade, who is the primary researcher on rubrics referenced in the literature. Dr. Andrade (personal
communication, October 1, 2006) confirmed that very few empirical studies have been conducted to test the claims that rubrics can improve academic achievement. The literature provides rationale for a number of possible relationships, like self-efficacy, worth exploring with respect to the claims of rubrics. Dr. Andrade is currently preparing a manuscript based on recent research on the effects of rubrics on self-efficacy for writing (Andrade, 2006a), lending credible support for the purpose of this study and recommendations for future research.

1. Future research should continue to explore the relationship of analytic trait, task-specific rubrics with self-efficacy for low efficacious students learning a comprehensive task. It is suggested the sample size be large enough that participants can be divided into high and low self-efficacy per treatment for comparison of main treatment effects and randomly assigned to conditions.

2. Future research should use subjects of different ages and abilities who have not formed high levels of skill and self-efficacy for a task. This change in methodology is consistent with the characteristics of participants in previous research on self-efficacy (Pajares & Miller, 1995; Schunk, 1991, 2003).

3. This study developed a new instrument to measure self-efficacy for writing a lesson plan and found it to have a moderate correlation between pre-task and post-task measures. Future research may be able to improve the correlation found in this study and further investigate whether the instrument correlates to any other established measurement instrument.

4. Future research should continue to investigate the impact on student achievement of using analytic trait, task-specific rubrics for a multi-
dimensional task that is challenging, but still within the zone of proximal development for the participant.

5. Future research should explore which specific features of analytic trait, task-specific rubrics may have impact on self-efficacy and achievement.

6. This study used brief explanation to inform participants of the benefits of rubric use. It is highly recommended to use a longer period of instruction with generative activities and practice using the rubric. It is suggested the study should be conducted over a longer period of time, such as one or two semesters, to better track the transfer of knowledge and skill to other tasks for which rubrics are used.

7. Future research should investigate which teaching strategies for analytic trait, task-specific rubrics result in the biggest impact. Schunk suggests strategy training can cultivate self-efficacy for learning when the learner understands and can effectively apply the strategy (1987b).

8. This study used a survey to collect post-task data on participant responses to the treatment and perceptions of the usefulness of the rubric. It may prove beneficial to revise the survey to better ascertain participants' actual use of the rubric and perceived challenge of the task. Following the study, it may be worthwhile to conduct a follow-up focus group with selected participants to uncover more dimensions of their experience using the rubric while accomplishing the task, that may otherwise not be clear from the survey responses.

9. See Appendix J for an alternative research study based on the premises of this study.
REFERENCES


APPENDIX A

Post Study Survey

Thank you for completing this survey. Mark your response as Indicated for each item.

1) What is your major or academic program?

2) Circle the period of years that matches your total classroom teaching experience:

<table>
<thead>
<tr>
<th></th>
<th>None</th>
<th>Less than 1</th>
<th>1-2</th>
<th>2-3</th>
<th>3-4</th>
<th>More than 4</th>
</tr>
</thead>
</table>

3) Rate your level of experience creating lesson plans by circling your response.

<table>
<thead>
<tr>
<th></th>
<th>Very little</th>
<th>Some</th>
<th>Quite a bit</th>
<th>A lot</th>
</tr>
</thead>
</table>

4) Thinking about the second lesson plan assignment you completed for this course, circle the total amount of time you estimate you spent working on creating the lesson plan.

<table>
<thead>
<tr>
<th></th>
<th>0 - 1 hours</th>
<th>1 - 3 hours</th>
<th>3 - 5 hours</th>
<th>5 - 7 hours</th>
<th>More than 7 hours</th>
</tr>
</thead>
</table>

5) Circle your response to the statement, “In general, when I write a lesson plan for academic credit or as a course assignment, I know I can get a good grade.”

<table>
<thead>
<tr>
<th></th>
<th>Always</th>
<th>Almost always</th>
<th>About half the time</th>
<th>Almost never</th>
<th>Never</th>
</tr>
</thead>
</table>

6) Circle your response to the statement, “In the past, I have used scoring rubrics to evaluate my own work on assignments or projects.”

<table>
<thead>
<tr>
<th></th>
<th>Frequently</th>
<th>Often</th>
<th>Sometimes</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
</table>

7) Circle your response to the question, “When you teach, how often do you use a scoring rubric to evaluate or score your students academic work/projects?”

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
</table>

- If you were not given a rubric to use to create your lesson plan assignment, stop and place this survey in the provided envelope. Thank you for your time.
- If you were given a rubric, please continue answering items 8 and 9.

8) Indicate how useful you found the rubric overall while you were developing your lesson plan.

<table>
<thead>
<tr>
<th></th>
<th>Not at all useful</th>
<th>A little useful</th>
<th>Somewhat useful</th>
<th>Very useful</th>
<th>Extremely useful</th>
</tr>
</thead>
</table>

9) Using this rating scale:

<table>
<thead>
<tr>
<th></th>
<th>A = strongly agree</th>
<th>B = agree</th>
<th>C = undecided</th>
<th>D = disagree</th>
<th>E = strongly disagree</th>
</tr>
</thead>
</table>

Please indicate your level of agreement by circling the letter that best describes your response to the following items that complete this statement, “The rubric I was given...”

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Helped me by providing explanations for how to do the lesson plan before I began working on it</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Did not help me understand the requirements to earn the highest possible score on the lesson plan assignment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Helped me establish the goals for creating my lesson plan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Did not provide me understandable feedback on how well I was completing the lesson plan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Did not save me time in completing the assignment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Helped me by providing information I could use to self-evaluate my work in creating the lesson plan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(05.10.08) v3
INSTRUCTIONS:

Use a black or blue ink pen.

Circle the number that corresponds to the first response that you think of after you read the items in this survey as honestly and genuinely as possible.

There are no right or wrong answers.

START HERE...

Using the rating scale shown below, circle the number that best describes how true the following statement is for you.

I am confident I will get an “A” grade on the next lesson plan assignment in this course.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not sure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Very sure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pretty sure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Somewhat sure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Not sure</td>
</tr>
</tbody>
</table>

When you have circled the number of your choice, turn the page and continue.
INSTRUCTIONS: Read the one question below, then for each item in the list use the rating scale shown and circle the number that best represents your response.

**How confident are you that you can successfully perform each of the items listed below?**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Not sure</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td><strong>Somewhat sure</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td><strong>Pretty sure</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td><strong>Very sure</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

**Read these items:**

1) Develop a lesson plan on a topic of your choice

2) Teach a topic using a lesson plan you developed

3) Write the purpose of a lesson

4) Write instructional objectives that help students accomplish the goal of the lesson

5) Write matching assessment items for each instructional objective

6) Integrate technology into the lesson for either the instructor or students to use

7) Select appropriate media, materials, and equipment for the lesson presentation

8) Create student activities that will show students have learned a particular objective

9) Describe specific knowledge, skills, or abilities that students should have after the lesson

10) Develop teaching strategies or activities that match the lesson and students’ learning style

11) Create a lesson plan using a word processor (i.e., Microsoft Word, AppleWorks)

12) Successfully complete an open-ended, complex course assignment *without using* a scoring rubric

13) Successfully complete an open-ended, complex course assignment *using* a scoring rubric

14) Meet your course instructor’s expectations for this lesson plan assignment

15) If your next lesson plan assignment were graded on a 100-point scale using the criteria described in items numbered 3 - 11 above, what score do you think you would earn?

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
</table>

When you have finished completing this survey, insert it into the provided envelope. Thank you!

(9.26.05)
APPENDIX C

Rubric pages 1 and 2
### Lesson Plan Rubric

**Criteria**

<table>
<thead>
<tr>
<th>Level</th>
<th>Exemplary (3)</th>
<th>Competent (2)</th>
<th>Developing (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong></td>
<td>Instructional purpose or “goal” is clearly stated in terms of what the learner is expected to be able to do upon completion of the lesson. Relevance of the lesson is clearly explained and meaningful to enable another instructor knowledgeable in this content area to deliver the lesson.</td>
<td>Instructional purpose or “goal” is stated in terms of what the learner is expected to be able to do upon completion of the lesson. Relevance of the lesson is explained well enough to enable another instructor knowledgeable in this content area to deliver the lesson.</td>
<td>Instructional purpose or “goal” is not clearly stated in terms of what the learner is expected to be able to do upon completion of the lesson. Relevance of the lesson is unclear and not explained in meaningful terms to enable another instructor knowledgeable in this content area to deliver the lesson.</td>
</tr>
<tr>
<td><strong>Objectives</strong> (3x)</td>
<td>Behavior, criteria, and conditions are clearly and concisely written (no unnecessary words). Learners can grasp a clear sense of what is expected of them and what they will achieve as a result of the lesson. Show a strong relationship between objectives and the instructional standards or purpose.</td>
<td>Behavior, criteria, and conditions are stated well enough to be understood. Learners might be able to grasp some sense of what is expected of them and what they will achieve as a result of the lesson. Show some relationship between objectives and the instructional standards or purpose.</td>
<td>Behavior, criteria, and conditions are not clearly written or contain unnecessary words. Learners cannot grasp a clear sense of what is expected of them or what they will achieve as a result of the lesson. Show a weak relationship between objectives and the instructional standards or purpose.</td>
</tr>
<tr>
<td><strong>Media, Equipment, Materials, Resources (Including people)</strong></td>
<td>All supporting materials needed for the instructor and learner to complete the lesson plan are clearly listed. Another instructor using this plan could clearly determine what materials need to be prepared in advance and exactly when to use them.</td>
<td>Most supporting materials needed for instructor and learner to complete lesson plan are listed. Another instructor using this plan could determine what materials need to be prepared in advance and when to use them.</td>
<td>Some or all supporting materials needed for the instructor and learner to complete the lesson plan are missing or not clearly listed. Another instructor using this plan could not easily determine what materials need to be prepared in advance or when to use them.</td>
</tr>
<tr>
<td><strong>Technology Used</strong> (2x)</td>
<td>Selection and application of computer-based technologies is appropriate for learning environment and outcomes. The technologies are appropriately applied to enhance learning.</td>
<td>Selection and application of computer-based technologies seem appropriate for learning environment and outcomes. The technologies could potentially enhance learning.</td>
<td>Selection and application of computer-based technologies are missing or not appropriate for learning environment and outcomes.</td>
</tr>
<tr>
<td><strong>Instructional Events</strong></td>
<td>Clearly articulates the sequence of events for how each objective will be taught from beginning to end. Each event has instructional strategies that include clear descriptions of the teacher activity and student activity.</td>
<td>Shows the sequence of events for how each objective will be taught from beginning to end. Each event has instructional strategies that include descriptions of the teacher activity and student activity.</td>
<td>Does not show the sequence of how each objective will be taught from beginning to end. Events do not have complete written instructional strategies that include descriptions of the teacher activity and student activity.</td>
</tr>
</tbody>
</table>

Continues...
<table>
<thead>
<tr>
<th>Criteria</th>
<th>Exemplary (3x)</th>
<th>Competent (2x)</th>
<th>Developing (1x)</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructional Strategies (3x)</td>
<td>Description of teacher activities provide clear instructions to guide the teacher. Description of student activities are appropriate for the target population and, if well-completed, will lead to learner achievement of the instructional objectives.</td>
<td>Description of teacher activities provide instructions to guide the teacher. Description of student activities are mostly appropriate for the target population and, if well-completed, should lead to learner achievement of the instructional objectives.</td>
<td>Description of teacher activities does not provide clear instructions to guide the teacher. Description of student activities are not appropriate for the target population or are not likely to lead to learner achievement of the instructional objectives.</td>
<td></td>
</tr>
<tr>
<td>Evaluation (3x)</td>
<td>Method for assessing student learning is clearly explained, matches the learning activities and objectives, and focuses on the essential elements of the lesson.</td>
<td>Method for assessing student learning is explained, mostly matches the learning activities and objectives and mostly focuses on the essential elements of the lesson.</td>
<td>Method for assessing student learning is not clearly explained, does not adequately match the learning activities and objectives or focuses on the wrong elements of the lesson.</td>
<td></td>
</tr>
<tr>
<td>Motivation (2x)</td>
<td>The overall purpose or &quot;big idea&quot; of the lesson is very likely to motivate and engage learners in meaningful learning. The instructor and student activities are engaging, creative, and innovative. They seem more &quot;fresh&quot; than &quot;stale.&quot;</td>
<td>The overall purpose or &quot;big idea&quot; of the lesson may motivate and engage learners in meaningful learning. The instructor and student activities are a mix of engaging, creative, or innovative, but seem standard for this content area.</td>
<td>The overall purpose or &quot;big idea&quot; of the lesson is not likely to motivate and engage learners in meaningful learning. The instructor and student activities are not engaging, creative, or innovative. They seem more &quot;stale&quot; than &quot;fresh.&quot;</td>
<td></td>
</tr>
<tr>
<td>Organization and Presentation</td>
<td>Lesson plan is printed, well organized and presented in a way that another instructor could easily read and understand. Organizes the content so that student understanding of new information can be enhanced.</td>
<td>Lesson plan is printed, organized and presented in a way that another instructor could read and understand. Organizes the content so that students can mostly understand the new information.</td>
<td>Lesson plan is not well organized or presented in a way that another instructor could easily read and understand. Organization of content interferes with students' ability to understand the new information.</td>
<td></td>
</tr>
</tbody>
</table>

SCORING: To calculate the score for each criterion, multiply the criterion weight factor* (1x, 2x, or 3x) by the performance level (exemplary = 3, competent = 2, developing = 1). Write the number in the Score column. The Total Score is the sum of the criteria scores. Maximum score = 51.

*Assume a weight factor of 1x ("one times") for criterion, unless otherwise indicated by 2x ("two times") or 3x ("three times").

**COMMENTS**
APPENDIX D

Rubric Presentation and Script
A Lesson on Rubrics

How can you use rubrics to improve, as well as judge, your own performance?

[Note: DO NOT give students the rubric until Slide #7, to prevent distracting them from paying attention to the presentation]

HOW TO READ THIS SCRIPT:
The words you speak are written on this notes area. I did not use quotation marks. The actions you take are in [brackets]. For some slides, I ask that you read the content of the slide. Otherwise, you may choose to read the content of the slide or not.

You now have your assignment to create your second lesson plan in this course.

As part of Prof. McGriff's research study on lesson plan instruction, he has prepared this presentation.

Let's read the objective for the lesson on the next slide.
Objective of the Lesson

- Students will be able to use the lesson plan rubric to improve and judge their own performance on the lesson plan assignment.

[Read this slide to the students]

For the research study, Prof. McGriff has prepared a rubric that will be used to score your lesson plans.

More important, you will be given the rubric today so you can use it to evaluate your lesson plan while you are developing it.

[next slide]
Rubric: Definition

- A format for organizing criteria, score points, and descriptions of good performance for judging the quality of a performance or product.

Let's be clear about the definition of a rubric. Take a minute to read this slide.

[Give 1 minute to read slide]

[Ask students for consensus on definition] 

*Does this definition match or closely fit your understanding of a rubric?*

[If students agree, go to next slide]

[If students disagree or offer alternative definitions, allow brief discussion, then state...]

Prof. McGriff derived this definition from a few sources and has based the remaining presentation on it.
Rubric

- Rubrics are powerful instructional tools for improving the very achievement that is also being assessed.
- Moving beyond using a rubric as an assessment tool and adding the instructional component.
- Teach students the criteria for quality and how to apply them to their own work to make it better.

In practice, rubrics could be more than an assessment tool used after the student has completed the project. It can be used by the student while developing the project to help them understand the standards of performance.

If a rubric is given with instruction on how to use it, the chances are very good that the student will be able to achieve a better performance on the project (e.g., higher score, improved quality, better match to standards).

If students are instructed in how to use the criteria in the rubric to assess the quality of their own project, then they can compare their own performance to the criteria. This is like doing a self-assessment or self-evaluation of the project while it is being completed.

In fact, that is the objective of this lesson on rubrics. If you understand how to use Prof. McGriff’s rubric to create your lesson plans and then you actually use the rubric to help you create your lesson plan, then you are likely to score higher than if you didn’t use the rubric.
Rubric: Benefits

When students know the performance criteria in advance, then there is...

- Shared vocabulary
- Clearer understanding of the important dimensions of quality performance = clear goals for student work

Let’s take a look at a few potential benefits of using the rubric.

[Allow students 1 minute to read slide]

[next slide]
Rubric: Benefits

- Opportunity to self-assess and improve their work along the way
- Enhance the quality of student learning and performance, not simply evaluate it
- Reinforces connection between standards-based education and performance criteria
  - Helps define standards, provides clear and consistent targets

More benefits include...

[allow students 1 minute to read slide]

The key benefits presented in this slide can be summarized as:
- Self-assessment tool
- Improve student performance
- Support standards-based education

[next slide]
Rubric

- What does it take to “know” and understand the performance criteria in order to improve achievement?

[allow students to read slide while you give them the rubric]

Let’s take a look at the performance criteria in the rubric.

[next slide]
Rubric

**Definition**
- a format for organizing criteria, score points, and descriptions of good performance for judging the quality of a performance or product

Remember our definition...

[allow students to read slide]
[allow students to read slide]

Take a look at the performance criteria in the rubric on page 1...it's that column on the left side.

These are the performance criteria that will used to judge the quality of your rubric.

[next slide]
Scale or Performance Levels

Labels that differentiate performance levels

- **Exemplary**
  - highest level of achievement

- **Competent**
  - the middle level is balanced between the high and low

- **Developing**
  - lowest level of achievement

Look at the labels across the top of the table...

These labels differentiate the performance levels by describing three different degrees of quality for each criteria.

[Ask rhetorically or “tongue-in-cheek”...] So, which performance level would you choose for your lesson plan? (-:

[next slide]
Descriptors

- Descriptions of what the criteria looks like at each performance level with statements of degree of success or achievement.
  - For example, read the descriptor for exemplary "Technology Used" criteria:
    - Selection and application of computer-based technologies is appropriate for learning environment and outcomes. The technologies are appropriately applied to enhance learning.

Take a look at the descriptors for each criteria and performance level.

[give students a moment to locate a descriptor, any descriptor]

Note the differing degrees of success, quality, or performance between the different levels (exemplary, competent, developing)

[next slide]
Scoring

- The value of the performance levels are shown with a number in parentheses.
  - 3 = exemplary, 2 = competent, 1 = developing
- The criteria are weighted to reflect their relative importance (2x, 3x).
- The criteria score is the performance level times the criteria weight.
- The total score is the sum of all criteria scores.
  - Total points (maximum score) = 51

A rubric is still a scoring tool so it also has a point value associated with the different levels of performance for each criterion.

The criteria are weighted differently to show the importance of them to the lesson plan. The different weights are either two times or three times the performance level.

For example:
1. The criterion “Objectives” is weighted to be three times the performance level value. So an exemplary score for objectives is 9 points (3 x 3).
2. The criterion “Technology used” is weighted two times the performance level value. So, an exemplary score is 6 points (2 x 3)

[next slide]
Using a Rubric

Using the rubric as an instructional tool

- Judging your performance (self-assessment) is the KEY to successfully using the rubric as a guide to completing the lesson plan assignment
- Review your lesson plan assignment before you begin to define all the work you need to complete

Now that you have the rubric, you may ask yourselves “how should I use it?”

Look at the rubric as an instructional tool to help you assess how well you are developing your lesson plan.

[allow students to finish reading slide]

[next slide]
When you are ready to begin your lesson plan...follow these basic steps.

You probably already do this type of process when starting an assignment.

[allow students to read slide or read the bullet points out loud]

[next slide]
Using a Rubric

*Using the rubric as an instructional tool*

- **Frequently review** and check your progress against the rubric criteria as you progress.
- **Check off** the criteria you believe you have accomplished for your lesson plan.

While you are in the middle of your lesson plan, come back to the rubric again and again

[allow students to read slide or read it out loud]

[next slide]
Using a Rubric

Using the rubric as an instructional tool

- When you think you are done with the lesson plan, critically analyze it against the criteria for the performance level you chose.
- Revise your lesson plan as needed to reach the quality level described for each criteria.

Finally, after you’ve finished the lesson plan, use the rubric for one more “double-check.”

If you find some weak areas, revise them to better match the performance level.

[allow students to read the slide or read it out loud]

[Answer questions as best you can without giving away the key point of the study. Students should feel they have received a challenge and a tool to help them meet the challenge.]

[next slide]
Bibliography


As you use the rubric to complete your lesson plans, do the best you can to interpret Prof. McGriff's rubric. Challenge yourself to see how well you can develop a lesson plan using the criteria of the rubric.

Good luck!

If you would like a copy of this presentation or the bibliography, Prof. McGriff will be glad to email it to you after you have completed your lesson plan. Send him an email with your request. Look at the bottom left footer of this slide to find his email address. His email is also printed on the informed consent form you received a copy of a short time ago.

Smogriff@email.ajsu.edu

©2005 Steven J. McGriff
APPENDIX E  IRB APPROVALS

1. IRB Approval from the Pennsylvania State University

2. IRB Approval from San José State University
Date: October 18, 2005

From: Tracie L. Kahler, IRB Administrator

To: Steven J. McGriff

Subject: Results of Review of Proposal – Exemption (IRB #21861)
Approval Expiration Date: October 17, 2006
Lesson Plan Instruction

The Office for Research Protections (ORP) has reviewed and approved your application for the use of human participants in your research. By accepting this decision, you agree to obtain prior approval from the ORP for any changes to your study. Unanticipated participant events that are encountered during the conduct of this research must be reported in a timely fashion.

If this study will extend beyond the above noted approval expiration date, the principal investigator must submit a completed Continuing Progress Report to the ORP to request renewed approval for this research.

On behalf of the ORP and the University, thank you for your efforts to conduct research in compliance with the federal regulations that have been established for the protection of human participants.

TLK/mhe
cc: Kyle L. Pack

Please Note: The ORP encourages you to subscribe to the ORP listserv for protocol and research-related information. Send a blank email to L-ORP-Research-L-subscribe-request@lists.psu.edu
To: Steven McGriff  
Instructional Tech., 0076  
Sweeney Hall 423

From: Pam Stacks, AVP  
Graduate Studies & Research

Date: June 27, 2005

The Human Subjects-Institutional Review Board has approved your request to use human subjects in the study entitled:

"Lesson Plan Instruction."

This approval is contingent upon the subjects participating in your research project being appropriately protected from risk. This includes the protection of the anonymity of the subjects' identity when they participate in your research project, and with regard to all data that may be collected from the subjects. The approval includes continued monitoring of your research by the Board to assure that the subjects are being adequately and properly protected from such risks. If at any time a subject becomes injured or complains of injury, you must notify Pam Stacks, Ph.D. immediately. Injury includes but is not limited to bodily harm, psychological trauma, and release of potentially damaging personal information. This approval for the human subjects portion of your project is in effect for one year, and data collection beyond June 27, 2006 requires an extension request.

Please also be advised that all subjects need to be fully informed and aware that their participation in your research project is voluntary, and that he or she may withdraw from the project at any time. Further, a subject's participation, refusal to participate, or withdrawal will not affect any services that the subject is receiving or will receive at the institution in which the research is being conducted.

If you have any questions, please contact me at (408) 924-2480.
APPENDIX F

Informed Consent Form
AGREEMENT TO PARTICIPATE IN RESEARCH

Responsible Investigator: Steven J. McGriff
Assistant Professor, Instructional Technology
San Jose State University
One Washington Square (Sweeney Hall 423)
San Jose, CA 95192-0076
Email: smcgriff@email.sjsu.edu
Phone: (408) 952-7198

Title of Protocol: Lesson Plan Instruction

Special Note: The responsible investigator is a doctoral candidate at The Pennsylvania State University, University Park, PA, working under the advisig of Dr. Kyle Peck, Professor, Instructional Systems. College of Education. This research study is being conducted for a doctoral dissertation.

1. Purpose of the Study: This study in which you can volunteer to participate is part of research intended to assess education students' abilities to create a lesson plan. The results of the study could lead to improved methods of instruction for learners and teachers.

2. Procedures: You will be asked to do the following:
   a. Allow the investigator to collect copies of your lesson plans completed as part of your normal course assignments in this semester;
   b. Complete a survey in class to evaluate your beliefs about lesson plans (16 items, using a Likert-type rating scale, approximately 5 minutes);
   c. OPTIONAL: Listen to a 20-minute presentation about lesson plans as homework before session 2;
   d. Create a lesson plan on your own time (duration established by your instructor);
   e. Complete a post-research survey in class to evaluate your beliefs about lesson plans (16 items, Likert-type rating scale, approximately 5 minutes)
   f. Complete a questionnaire in class about your experiences creating your lesson plan (1 page, approximately 5 minutes)

3. Discomforts and Risks: There are no physical, psychological, or emotional risks in participating in this research beyond those experienced in your normal classroom and educational environments.

4. Benefits:
   - The benefits to participants include exposure to instructional, learning, and assessment strategies that could be applied to the teaching of any subject.
   - The potential benefits to society include knowledge of how teachers can improve their students' academic achievement by using a prescribed combination of instructional and assessment strategies.

5. Alternative Procedure: If you choose not to participate or cannot complete the study, your course instructor will provide an alternative assignment of equivalent value.

Initials
6. **Confidentiality:** Your participation in this research study is confidential. Only the responsible investigator will be able to match your identity with the information you provide. In the event of publication of this research, no personally identifying information will be disclosed. To make sure your participation is confidential, your identifying information will be erased and replaced with a code number on the instructional materials and research instruments. Only the researcher can match names with code numbers. Your data will be secured in a locked filing cabinet in a secured office.

7. **Compensation:** The compensation for participation in this study is academic credit in your class, as arranged with your course instructor.

8. **Right to Ask Questions:** You have the right to ask questions at any time during the course of this research study and can be addressed to Steven J. McGriff at (408) 924-3654 or by email (smegriff@ccmail.sju.edu). Complaints about the research may be presented to Dr. Roberta Barba, Chair, Instructional Technology Department, College of Education, at (408) 924-3613. Questions about research subjects’ rights, or research related injury may be presented to Pamela Stacks, Ph.D., Associate Vice President, Graduate Studies and Research, at (408) 924-2490.

9. **Right to Not Participate:** No service of any kind, to which you are otherwise entitled, will be lost or jeopardized if you choose to “not participate” in the study.

10. **Voluntary Participation:** Your consent is being given voluntarily. You may refuse to participate in the entire study or in any part of the study at any time. If you decide to participate in the study, you are free to withdraw at any time without any negative effect on your relations with your course instructor, San Jose State University, or The Pennsylvania State University.

11. **Signature:** You must be 18 years of age or older to consent to take part in this research study. At the time you sign this consent form, you will receive a copy for your records, signed and dated by the investigator.

- The signature of a subject on this document and their initials on the first page indicates the subject has read this informed consent form and agrees to participate in the research study.

- The signature of a researcher on this document indicates agreement to include the above named subject in the research study and attestation that the subject has been fully informed of his or her rights.

PRINT Participant First and Last Name

<table>
<thead>
<tr>
<th>Participant Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Steven J. McGriff, responsible investigator

<table>
<thead>
<tr>
<th>Date</th>
</tr>
</thead>
</table>
APPENDIX G

Instructor C Lesson Plan Assignment
Purpose: Demonstrate your ability to plan, instruct, assess and reflect on a single lesson designed to meet the learning needs of your students.

Overview -- What to include:
A. Planning
   • final lesson plan or planning matrix (addressing all planning “big ideas”)
   • planning conversation form (left side on your own, right side with CT)
   • rationale for lesson design (see prompts below)

B. Instructing
   • videotape of lesson
   • reflection form (with CT after watching videotape)

C. Assessing Student Learning
   • copy of assessment tool or prompt
   • three samples of student work (at least one from English learner)
   • assessment commentary (see prompts below)

D. Reflection
   • reflection commentary (see prompts below)

Part A – Planning
Work in collaboration with your cooperating teacher to identify a focus for the lesson. Write a draft of your lesson plan and get feedback from your CT on your ideas. Use the planning conversation form to share your lesson and record your CT’s feedback. Revise your lesson plan and include the final draft in your coaching cycle materials. Write a rationale for the design of your lesson addressing the following prompts:

1. Why is the content of this lesson important for your particular students to learn? (consider its importance apart from being in the school curriculum or academic standards for this grade level)

2. Explain why the instructional tasks or strategies you are using will support the learning of YOUR students (consider any relevant theories as well as who your students are, what they bring, and the needs of ELs, GATE students, or those with IEPs).

3. How do you plan to address and support the needs of students who might find the language demands of your lesson challenging?

4. How will you assess what your students have learned related to the specific standards and objectives of the lesson?
B. Instructing
Teach and videotape your lesson. Watch the tape with your cooperating teacher and complete the reflection form.

C. Assessing Student Learning
Select three student work samples which together represent what students generally understood and what a number of students were still struggling to understand in this lesson. At least one of these students should be an English Learner. Address the following prompts in your commentary:

5. Identify the specific standards/objectives for this student assessment; and describe the corresponding evaluative criteria you used to determine student learning using this assessment.

6. In relation to your evaluative criteria, discuss what the three students appear to understand well, and any misunderstandings, confusions, or needs (including a need for greater challenge) that were apparent. What can you conclude about their learning of the desired outcomes of this lesson? Cite evidence to support your analysis from the work samples you selected.

7. Based on students’ performance on this assessment, describe the next steps for instruction for the class, and any individualized next steps for the students whose individual learning you analyzed. These next steps may include feedback to students, a specific instructional activity, or other forms of re-teaching to support or extend student learning. In your description, be sure to explain how these specific next steps follow from your analysis of student performance.

D. Final Reflection
Use information from the reflection form completed with your cooperating teacher and your analysis of student work to address the following prompts:

8. What is working in this lesson? For whom? Why?
9. What is not working? For whom? Why?
10. How does this teaching/reflection process inform what you plan to do next?
11. How does your assessment of student work inform what you plan to do next?
12. Looking back on this lesson, what would you do differently if you could teach this lesson again to the same group of students? (Consider changes to planning, teaching or assessment.) How would these changes improve learning for your students, especially those with particular needs?
APPENDIX H

Instruction Guide for Raters

The following 7 pages comprise the instructor guide used to train lesson plan raters.
### Unit Title
Training for Inter-rater Reliability

| Grade Level: | Professional Teacher |

### Lesson Title
How to Rate a Lesson Plan

### Purpose
Raters will learn the purpose of and procedure for evaluating lesson plans written by participants in a research study. In addition, raters will use the lesson plan rubric to effectively evaluate the level of subjects’ achievement for developing a lesson plan.

### Objectives
1. Describe purpose of the research project and lesson plan evaluation process
2. Explain inter-rater reliability
3. Identify parts of the rubric
4. Express performance levels in own terms
5. Record a score for each performance criteria
6. Describe the procedure for evaluating the lesson plan
7. Execute the procedure for managing the evaluation process
8. Record rater’s notes on evaluation process

### List of Media, Equipment, Materials, Resources
- Laptop computer
- Handout: Procedure Description (TBD)
- Handout: Inter-rater reliability info (TBD)
- Sample lesson plans (one for each performance level) (TBD)
- Rater-Score sheets (TBD)
- PowerPoint slideshow (lessonplan.ppt)
- Lesson Plan Assignment sheet
- PowerPoint (rubricLesson.ppt)
- Handout: A Lesson on Rubrics

### Lesson Procedure
This training is organized in seminar format, which allows for team-building, focused instruction, and practice that maximizes trainer-participant interaction. Proceed at the fastest pace allowed with frequent checks for understanding and paraphrasing.
### Objective 1:
Given instruction, the rater will be able to accurately describe the purpose of the research project and lesson plan evaluation process.

<table>
<thead>
<tr>
<th>Instructional Event</th>
<th>Learning Outcome</th>
<th>Teacher Activity</th>
<th>Student Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gain attention</td>
<td>Generate students motivation and interest in the topic and/or objective</td>
<td>❑ Express appreciation</td>
<td>Take notes, ask questions, seek clarification</td>
</tr>
<tr>
<td>Main presentation</td>
<td>Students receive new instructional material in a sequence that builds knowledge or skill</td>
<td>❑ Discuss and explain:</td>
<td>Review materials:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>❑ Goal of research project</td>
<td>Lesson Plan Assignment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>❑ Lesson plan assignment</td>
<td>Assignment sheet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>❑ Review lesson plan instructional materials</td>
<td>Review rubric lesson materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>❑ Review rubric lesson materials</td>
<td>Lesson Plan evaluation process description</td>
</tr>
<tr>
<td></td>
<td></td>
<td>❑ Role of the rater in research to perform consistently</td>
<td>Role of the rater in research to perform consistently</td>
</tr>
<tr>
<td>Elicit student</td>
<td>Students activate cognitive processing to help integrate new information or develop new skills, abilities</td>
<td>Seek feedback from participants. Note their activity and interaction with materials.</td>
<td>Active discussion and response.</td>
</tr>
<tr>
<td>performance</td>
<td></td>
<td>Ask participants to describe the purpose of the research project and lesson plan evaluation process.</td>
<td></td>
</tr>
<tr>
<td>Provide feedback on</td>
<td>Students synthesize lesson content and receive feedback on progress</td>
<td>Answer questions and encourage divergent thinking; engaging participant as co-researcher on this project.</td>
<td></td>
</tr>
<tr>
<td>performance</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Steven McGriff
### Objective 2:
Given instruction, the participant will be able to correctly explain the purpose of inter-rater reliability and the role of the rater

<table>
<thead>
<tr>
<th>Instructional Event</th>
<th>Learning Outcome</th>
<th>Teacher Activity</th>
<th>Student Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main presentation</td>
<td></td>
<td>Discuss:</td>
<td>Review materials:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>❑ Inter-rater reliability concept</td>
<td>❑ Inter-rater reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>❑ procedures/methods</td>
<td>❑ Procedure Description</td>
</tr>
<tr>
<td></td>
<td></td>
<td>❑ Objectivity, consistency</td>
<td></td>
</tr>
<tr>
<td>Elicit student performance</td>
<td></td>
<td>Seek feedback from participants.</td>
<td>Active discussion and response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Note their activity and interaction with materials.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ask participants to describe the role of the rater and need for objectivity and</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>consistency.</td>
<td></td>
</tr>
<tr>
<td>Provide feedback on performance</td>
<td></td>
<td>Answer questions and encourage divergent thinking; engaging participant as co-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>researcher on this project.</td>
<td></td>
</tr>
<tr>
<td>Closure</td>
<td></td>
<td>Summarize key points of inter-rater role and process</td>
<td></td>
</tr>
</tbody>
</table>

### Objective 3:
Given the rubric, the participant will be able to correctly identify the parts of the rubric (11 performance criteria, 3 performance levels, score points).

<table>
<thead>
<tr>
<th>Instructional Event</th>
<th>Learning Outcome</th>
<th>Teacher Activity</th>
<th>Student Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main presentation</td>
<td></td>
<td>Discuss and review lesson plan rubric using:</td>
<td>Review materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PowerPoint (rubricLesson.ppt)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Handout: A Lesson on Rubrics</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cover each part of the rubric with explanation and discussion</td>
<td></td>
</tr>
<tr>
<td>Elicit student performance</td>
<td></td>
<td>Seek feedback from participants.</td>
<td>Active discussion and response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Note their activity and interaction with materials.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ask participants to describe the parts of the rubric in their own terms</td>
<td></td>
</tr>
<tr>
<td>Provide feedback on performance</td>
<td></td>
<td>Answer questions and encourage active use of the rubric as a tool for assessment</td>
<td></td>
</tr>
<tr>
<td>Closure</td>
<td></td>
<td>Summarize key categories of rubric and central role in this evaluation process</td>
<td></td>
</tr>
</tbody>
</table>

Steven McGriff
### Objective 4
Given the rubric, the participant will be able to express the degrees of difference between the 3 performance levels in their own terms.

<table>
<thead>
<tr>
<th>Instructional Event</th>
<th>Learning Outcome</th>
<th>Teacher Activity</th>
<th>Student Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main presentation</strong></td>
<td>Focus on the three different performance levels: Exemplary, Competent, Developing</td>
<td>Handout: Rubric</td>
<td>Review materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Handout: A Lesson on Rubrics</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample lesson plans</td>
<td></td>
</tr>
<tr>
<td><strong>Elicit student performance</strong></td>
<td>Ask participants read through all the criteria and performance level descriptions.</td>
<td>Ask participants to paraphrase the differences between performance levels</td>
<td>Ask participants to review the sample lesson plans and match it with the rubric descriptions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Review examples of lesson plans that show different performance levels</td>
<td></td>
</tr>
<tr>
<td><strong>Provide feedback on performance</strong></td>
<td>Answer questions and encourage consistent use of levels, offer tips on evaluating differences</td>
<td></td>
<td>Ask questions</td>
</tr>
<tr>
<td><strong>Closure</strong></td>
<td>Summarize differences between levels</td>
<td></td>
<td>Discussion</td>
</tr>
</tbody>
</table>
Objective 5
Given the rubric and a real lesson plan, participants will be able to correctly record a score for each performance criteria and calculate a total score for the lesson plan.

<table>
<thead>
<tr>
<th>Instructional Event</th>
<th>Learning Outcome</th>
<th>Teacher Activity</th>
<th>Student Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main presentation</td>
<td></td>
<td>Walk through a real lesson plan from the study, demonstrating my scoring techniques</td>
<td>Review materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Focus on the three different performance levels and the score points associated with each: Exemplary = 3, Competent = 2, Developing = 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Handout: Rater Score Sheet</td>
<td></td>
</tr>
<tr>
<td>Elicit student performance</td>
<td></td>
<td>Demonstrate how to write score points in score column and total the score at the bottom of the last sheet.</td>
<td>Follow along and practice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Participants practice scoring the sample lesson plans</td>
<td></td>
</tr>
<tr>
<td>Provide feedback on</td>
<td></td>
<td>Answer questions and encourage consistent use of levels and score points</td>
<td></td>
</tr>
<tr>
<td>performance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closure</td>
<td></td>
<td>Summarize score point values</td>
<td>Discussion</td>
</tr>
</tbody>
</table>
### Objective 6
Given a set of scored sample lesson plans, the participant will be able to competently compare their scored lesson plan with the samples.

<table>
<thead>
<tr>
<th>Instructional Event</th>
<th>Learning Outcome</th>
<th>Teacher Activity</th>
<th>Student Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main presentation</td>
<td></td>
<td>Review procedure description and rater score sheet</td>
<td>Review materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Demonstrate how to compare participants’ newly scored lesson plan with models of good scoring shown on sample lesson plans</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample scored lesson plans</td>
<td></td>
</tr>
<tr>
<td>Elicit student performance</td>
<td></td>
<td>Oversee, coach, and guide participants practice of comparative review</td>
<td>Conduct comparative review</td>
</tr>
<tr>
<td>Provide feedback on</td>
<td></td>
<td>Answer questions and encourage familiarity with and consistent use of procedure</td>
<td>Exchange scored lesson plans with other participants and conduct review again</td>
</tr>
<tr>
<td>performance</td>
<td></td>
<td>Conduct formative reviews until participants demonstrate consistency</td>
<td></td>
</tr>
<tr>
<td>Closure</td>
<td></td>
<td>Summarize comparative review process and value to achieving consistent scoring</td>
<td></td>
</tr>
</tbody>
</table>
**Objective 7**
Given instruction, the participant will be able to competently execute the procedure for managing multiple lesson plan evaluations

<table>
<thead>
<tr>
<th>Instructional Event</th>
<th>Learning Outcome</th>
<th>Teacher Activity</th>
<th>Student Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main presentation</td>
<td></td>
<td>Handout: Procedure Description Sets of lesson plans from study</td>
<td>Review materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Review procedure description for handling all the lesson plans the rater has</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Describe how to store, check-out, and return lesson plans</td>
<td></td>
</tr>
<tr>
<td>Elicit student performance</td>
<td></td>
<td>Demonstrate how to execute procedure, step-by-step</td>
<td>Follow along and practice</td>
</tr>
<tr>
<td>Provide feedback on performance</td>
<td></td>
<td>Answer questions and encourage familiarity with and consistent use of procedure to handle lesson plans in the stack the rater has</td>
<td>Practice again</td>
</tr>
<tr>
<td>Closure</td>
<td></td>
<td>Summarize procedure</td>
<td></td>
</tr>
</tbody>
</table>

**Objective 8**
Given instruction, the participant will be able to consistently record their notes on the evaluation process for each lesson plan

<table>
<thead>
<tr>
<th>Instructional Event</th>
<th>Learning Outcome</th>
<th>Teacher Activity</th>
<th>Student Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main presentation</td>
<td></td>
<td>Demonstrate how to review a lesson plan and record notes on the Rater Score sheet for the researcher to interpret</td>
<td>Review materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Demonstrate how to complete the Rater Score sheet with consistent language, tone, and completeness</td>
<td></td>
</tr>
<tr>
<td>Elicit student performance</td>
<td></td>
<td>Observe raters while they practice</td>
<td>practice using the sample scored lesson plans</td>
</tr>
<tr>
<td>Provide feedback on performance</td>
<td></td>
<td>Answer questions and encourage consistent use of Rater Score sheets to handle lesson plans in the stack the rater has received</td>
<td></td>
</tr>
<tr>
<td>Closure</td>
<td></td>
<td>Summarize procedure</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX I

Overview of Research Procedure
APPENDIX J

Alternative Research Study

Explanation of this Alternative Research Study

This alternative research study is included here to show another method by which the dissertation study could be conducted. It is built upon the same basic problem statement and literature. The methodology differences include the audience, which was changed from teacher credential candidates to undergraduate chemistry students and the task has changed from writing a lesson plan to writing a lab report. Most importantly, this alternative study uses a pretest-posttest control group design (3 x 2 factorial) in order to better control for threats to internal validity.

Title

The Effect of an Analytic, Task Specific Rubric Implementation Strategy on the Self-efficacy of Introductory College Chemistry Students and Achievement for Writing a Lab Report

Purpose of Study

The main purpose of this research study is to investigate the effect of three strategies for implementation of analytic trait, task-specific rubrics—no rubric, rubric given without instruction, and rubric given with instruction—on introductory college chemistry student’s self-efficacy for writing a lab report.

Results of this study could be applied to the training of chemistry faculty. In particular, criteria for an exemplary lab report would be suggested, as well as an instructional strategy for implementing the use of a rubric for guiding students’ development and instructors’ evaluation of a lab report.
Research Questions

1. Does the strategy for implementation of rubric use lead to differences in post-task self-efficacy for writing a lab report among three treatment groups?

2. Does the strategy for implementation of rubric use lead to differences in post-task self-efficacy for writing a lab report between high and low self-efficacy participants?

3. Does the strategy for implementation of rubric use lead to differences in performance scores for writing a lab report among three treatment groups?

4. Does the strategy for implementation of rubric use lead to differences in performance scores for writing a lab report between high and low self-efficacy participants?

Literature Review

See Chapter 2 for the basis of the literature that threads together the relationships between self-efficacy, rubrics, and achievement on a task.

Research Design

Pretest-posttest control group design (3 X 2 factorial)

According to Gall and colleagues (2005), a true experimental design, like pretest-posttest control group, has all three essential characteristics of a good experiment: random assignment; administration of a treatment to an experimental group, an alternative treatment to a second group (and in this study, no treatment to a control group); and comparison of the groups’ performance on a post treatment measure. The factorial design is derived from 3 conditions with 2 levels of self-efficacy (high, low) in each.
Threats to internal validity

Stanley and Campbell (1963) identified eight types of extraneous variables that can affect the internal validity of the experiment. The following is a brief discussion of the affect of each threat on this study.

**History effect.** Given participants will be engaged in this study for four weeks in a semester, there exists a threat that other instruction from teachers could account for all or part of the participant’s achievement gain. Randomization of participants is the best solution for minimizing this effect.

**Maturation effect.** The psychological changes that occur in first and second year college students are not expected to be a threat during the relatively short time of the study. Randomization of participants is the best solution for minimizing this effect.

**Testing effect.** Given a pretest-posttest measurement for self-efficacy, this threat can be minimized by allowing the maximum time between administrations of the instrument. In addition, the pretest measure can be used as a covariate when analyzing the posttest scores in order to “wash-out” its influence. Regarding measuring achievement gain for writing a lab report, another assignment correlated to the lab report and completed before treatment or group assignment can be used as a covariate for the posttest achievement measure.

**Instrumentation effect.** Using the same self-efficacy instrument for the pretest and posttest minimizes this threat. The rubric used throughout the study would be the same.

**Statistical regression.** Randomization of participants is the best solution for minimizing this effect, which will happen in all pretest-posttest studies.

**Differential selection.** Randomization of participants is the best solution for minimizing this effect.
Selection-maturation interaction. Randomization of participants drawn from the same population is the best solution for minimizing this effect.

Attrition. This threat is very real for a study that lasts a full semester and is partially based on classroom participation. To minimize the effect, the control would participate in the same number of outside instructional sessions as the experimental conditions.

Generalizability

The results of the study can be generalized to the population of college students taking introductory chemistry classes in a U.S. educational institution with similar characteristics.

Sampling

The sample population is drawn from the undergraduates, namely freshmen, enrolled in a general education, introductory chemistry course (CHEM 001A) at a large, public, west coast university.

CHEM 001A: General Chemistry. Course topics include stoichiometry, reactions, atomic structure, periodicity, bonding, states of matter, energy changes, and solutions using organic and inorganic examples. Lab program complements lecture. Prerequisite: Proficiency in high school chemistry or Chem. 10 (with a grade of "C" or better; "C-" not accepted) or instructor consent; proficiency in high school algebra and eligibility for Math 19; eligibility for Engl 1A. Misc/Lab: Lecture 3 hours/lecture 1 hour/lab 3 hours. Units: 5.

Volunteers are invited to participate via a scripted presentation from the principal researcher in cooperation with the course instructors, who would offer extra credit. It is hoped that if the course instructor serves as an advocate for the study,
participants will be more likely to engage with full attention for the duration, thus minimizing attrition. The population of CHEM 001A students in a typical Fall semester is more than 650. The optimal sample size needed for this study is 300 participants randomly assigned to three conditions. With approximately 100 participants per condition, two levels of self-efficacy (high and low) can be created with approximately 40 participants in each level, eliminating about 10 participants in the mid-range in order to build greater difference between the levels. Participants would not be informed of their self-efficacy level nor treated any differently because of it.

**Variables**

| Table J.1 Alternate Research Study List of Variables |
|-----------------------------------------------|-----------------------------------------------|
| **Variable** | **Type** | **Measurement Instrument** | **Data Analysis Procedure** |
| Treatment PreSE | IV Control Rubric Rubric + instruction | MSLQ | ANOVA |
| PostSE | DV High Low | MSLQ | ANCOVA |
| Pretest Post Score | Covariate | Chemistry Test 1 Rubric for lab report | ANOVA ANCOVA |

**Methods of Data Collection**

**MSLQ (Motivated Strategies for Learning Questionnaire).** Self-efficacy (PreSE and PostSE) is measured using the MSLQ, a self-report instrument designed to assess college students’ motivational orientations and their use of different learning strategies for a college course. The motivation section consists of 31 items that assess, among other
things, self-efficacy. The learning strategy section includes 31 items regarding students’ use of different cognitive and metacognitive strategies. In addition, 19 items concerning student management of different resources is included in the learning strategies section. The self-efficacy subscale has a Cronbach’s Alpha reliability of .93 and a correlation to final grade of .41.

**Chemistry Test 1 (Pretest).** The first test completed in the course serves as a covariate for the Post Score. Reliability will be determined from previous administrations of the instrument. The type of extant evidence that will be gathered and analyzed in collaboration with the CHEM 001A faculty to establish validity includes: test content, internal structure, and relationship to other variables.

**Rubric for lab report (Post Score).** The rubric is a score sheet used to evaluate the quality of the written lab report. The score is the average of two scores as determined by two independent raters in blind-review process. A pilot study will be used to determine the reliability of the rubric and gather evidence of validity for rubric content, internal structure, and relationship to other variables.

**Ethics and Human Relations**

IRB approval is required for the study. All faculty collaborators will have completed mandatory training in human subjects. There are no perceived threats to participants beyond the normal occurrences of classroom instruction.

Access to the CHEM 001A setting will be facilitated through an established relationship with Dr. Resa Kelly, Assistant Professor of Chemistry, who asked the principal researcher to help the Chemistry Department develop a rubric for evaluating lab reports. Leveraging the collaborative relationship and the buy-in of some Chemistry faculty members, the CHEM 001A course instructors will be recruited to participate as a
service to their department. Participants in the study will be recruited in collaboration with the course instructors, who will provide incentives, such as extra credit for complete participation.

Timeline

One or two semesters before the main study, develop reliable rubric and pretest instruments and establish validity in collaboration with CHEM 001A instructors.

- Semester week 1: Recruit volunteer participants
- Semester week 2: Administer MSLQ (PreSE), then Chemistry Test 1 (Pretest)
- Semester week 3: Begin first of 3 weekly extracurricular instruction sessions on using a rubric to write a lab report.
- Semester week 6: Collect lab reports completed after last instruction session.
- Semester week 12: Administer MSLQ (PostSE).

End of study procedure
STEVEN J. MCGRIFF
Assistant Professor, Department of Instructional Technology
College of Education, San José State University
One Washington Square, San Jose, California 95192-0076
smcgriff@email.sjsu.edu www.sjsu.edu/depts/it/mcgriff

EDUCATION
The Pennsylvania State University, University Park, PA —
Ph.D., Instructional Systems, College of Education
Supporting Field: Leadership in Higher Education
Award: Academic Computing Fellowship
San José State University, College of Education, San José, CA —
M.A, Education: Instructional Technology
Stanford University, Stanford, CA —
B.A., Studio Art: Art Design

EXPERIENCE
Assistant Professor, Instructional Technology August 2005 – present
College of Education, San José State University, San José, CA
http://www.sjsu.edu/depts/it/mcgriff
Lecturer, Instructional Technology August 2002 – August 2005
College of Education, San José State University, San José, CA
Lecturer, Universidad Don Bosco—San José State University
San Salvador, El Salvador
Instructional Design Consultant Intern. June – August 2001
Innovative Learning, Andersen, St. Charles, IL
Manager, Academic Technology Lab March 1998 – August 1999
Stanford University, Stanford, CA

PROFESSIONAL ASSOCIATIONS
Association for Educational Communications and Technology (AECT)
Minorities in Media (AECT National Affiliate)
American Education Research Association (AERA)
International Society for Technology in Education (ISTE)
International Visual Literacy Association (IVLA)
Computer Using Educators (CUE)
Association for State Technology-Using Teacher Educators (ASTUTE)
iCUE-Silicon Valley (local chapter of CUE)