COMPARISON OF EFFECTS OF COGNITIVE LEVEL AND QUALITY WRITING ASSESSMENT (CLAQWA) RUBRIC ON FRESHMAN COLLEGE STUDENT WRITING

A Dissertation

Presented to

The Faculty of the School of Education

Liberty University

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

by

Irene Suzanne Penner

November 2010

Comparison of Effects of Cognitive Level and Quality Writing Assessment (CLAQWA) Rubric on Freshman College Student Writing

by Irene Suzanne Penner

APPROVED:

COMMITTEE CHAIR	Kathie C. Morgan, Ed.D.
COMMITTEE MEMBERS	William G. Gribbin, Ph.D.
	John R. Duryea, Ed.D.
CHAIR, GRADUATE STUDIES	Scott B. Watson, Ph.D

ABSTRACT

Penner, Irene Suzanne. COMPARISON OF EFFECTS OF COGNITIVE LEVEL AND QUALITY WRITING ASSESSMENT (CLAQWA) RUBRIC ON FRESHMAN COLLEGE STUDENT WRITING. (Under the direction of Dr. Kathie C. Morgan) School of Education, November, 2010.

The study investigated the effects of the Cognitive Level and Quality Writing Assessment (CLAQWA) rubric on the cognitive skill and writing skill growth of college freshmen. The participants (n = 107) were enrolled in a composition course at a Midwestern state university. The nonequivalent control group design used quantitative analysis with selected criteria from the CLAQWA rubric as measurements. Two independent raters graded the essays, and results confirmed a statistically significant correlation of grades on both sets of essays. Results from both raters confirmed no statistically significant differences on either type of skill score between the experimental or control group for the final essay. Results suggest that although a specific rubric enhances the learning environment, a specific rubric does not define the learning environment. Results also demonstrated a statistically significant difference between the female and male groups for the diagnostic essays graded by rater one; however, there was no statistically significant difference between male and female groups on the final essay as graded by rater two. Results indicated that the measurement of student outcomes, mandated by recent legislative efforts, may be accomplished through the use of a rubric, but at the same time, a specific rubric may not be a universal answer.

Keywords: CLAQWA, cognitive writing model, cognition, prescriptive writing, descriptive writing, writing skills, freshmen composition

Dedication

ñWhatever you do, work at it with all your heart, as working for the Lord, not for men, ... It is the Lord Christ you are serving. Colossians 3:23-24

ñOh, that you would bless me . . . and let your hand be with me . . . . And God granted his request. I Chronicles 4:10

I praise God for seeing me through the last eight years of working on this degree. I pray that I may be able to pass it on.

Acknowledgements

My successful completion of the course work, research, and writing would not have been possible without the ongoing support of my husband and children, and the libraries of Bridgewater State College, Cape Cod Community College, and Youngstown State University. I owe a special acknowledgement to Steven McDonald, who graciously donated his time and help to the statistical reports of this dissertation. I appreciate all effort from others toward the successful completion of this endeavor.

I would also like to thank: Dr. Kathie C. Morgan, my dissertation chair, who promptly and untiringly responded to many needs and questions; to Dr. William G. Gribbin, who talked me through the work; and to Dr. John R. Duryea, whose valuable comments helped me conclude the task. I have enjoyed working with all of you, and may God bless you as you move on to help others.

I have greatly appreciated the ongoing effort of Liberty University in upholding the Word of God. The perspective that comes from understanding and applying biblical principles to the arena of education is very refreshing and has helped me to understand how much God communicates his love to others through the world of teaching. As a result of earning this degree, I would like to teach full-time at a Christian college, where I may hope to inspire students in the same way I have been inspired.

Table of Contents

List of Tables

List of Figures

Chapter One: Introduction

Undergraduate institutions have recently been faced with accountability mandates from state and accreditation sources. These challenges need not perplex or daunt any departments or colleges, but challenges must be addressed. This dissertation describes a nonequivalent control group design that examined differences in outcomes between an experimental treatment and a control treatment in teaching freshmen writing courses. The study focused on the final/third assigned essay of the semester to learn whether using the experimental treatment resulted in more significant writing progress than using the control treatment. The first/diagnostic essay of the semester was compared to learn whether both groups began at the same level of competency. This study addressed an area about which many have legitimate concerns.

Background of Composition Studies

This research study examined the differences between the cognitive level and writing quality of college students enrolled in ten Writing I courses. The study was based primarily on the implementation of a specific rubric within the freshman college writing classroom at an urban, Midwestern state university. All freshmen have been required to take Writing I, and the course has typically expected students to show writing improvement as a result of completing the course. The final papers of the five experimental classes and the five control classes were compared using a specific rubric. This first chapter introduces the historical background to the theory behind the study, states the research problem, explains the professional significance of the research problem, and gives the operational definitions that pertain to this study.

Legislative. The three background factors which have influenced the research upon which this dissertation is based include: 1) the legislative background, 2) the professional background, and 3) the research-oriented background. Identifiable shifts in theories about education and

perceptions about learning (Bloom, 1956, 1984; Chickering & Gamson, 1987; Perry, 1999) have brought about specific trends and practices that in turn impact legislative mandates (*Action Plan*, 2006; *Boards*, 2006; *Executive*, 2007; *Four Pillars*, 2004; *U.S. Department*, 2007); legislation affects education at every level. An increasing concern with institutional accountability coupled with how students learn and with how curriculum meets the needs of students is now being examined at the college level.

At the beginning of the twenty-first century, a significant shift in education occurred when the No Child Left Behind Act (NCLB) emphasized academic accountability more than other guidelines for schools; and the Act underscored accountability for results rather than accountability for programming, as did previous acts of education. The nature of this legislation has put pressure on schools to prove that students have actually attained goals of academic improvement (*Four Pillars*, 2004). Former guidelines for education allowed schools to emphasize curriculum without overt emphasis on outcomes, but the academic climate has changed drastically.

Currently, the NCLB now emphasizes four guidelines by which schools can operate most effectively: ñStronger Accountability for Results. é More Freedom for States and Communities. é Proven Education Methods. é [and] More Choices for Parentsò (*Four Pillars*, 2004). Just as K-12 public schools are presently trying to close a gap between methods and outcomes, post-secondary institutions are now trying to address a gap between curriculum and outcomes. Corporate responsibility for results at the K-12 level appears to have initiated a similar interest in accountability, affordability, quality, and access at the college level.

One obvious feature of this directional shift in education has involved a more student-oriented approach to teaching instead of a curriculum-oriented or professor-oriented focus

(Chickering & Gamson, 1987; Perry, 1999). Such a shift makes sense at the college level, where students enter the higher level of education already in possession of certain definable cognitive skills and writing quality skills. The more student-oriented approach works well in freshman college writing courses because the larger purpose of the general education requirement courses (such as freshman writing) is to prepare undergraduates for the baccalaureate level of education. In other words, the nature of freshman writing courses implies a more comprehensive scope than the departmental designation might suggest. In view of that broader range, freshman writing courses can focus on cognition as a separate component from content.

Although the passage of the NCLB legislation has mandated student assessments, even the Educational Testing Service has observed that many educators do not fully understand them, nor do they use them correctly (Olson, 2005, Curriculum section, para. 6). According to Olson, part of the challenge has involved not knowing how to translate state and district standards into classroom practice. Because educators do not necessarily see the importance of these standards in daily lesson planning (Olson, 2005, Curriculum section, para. 8), textbooks that have included examples of how to incorporate the standards may be useful to educators when implementing recent legislation. And certainly, the purposes of testing need to be made clearer, especially to those who administer the assessments.

After the federal government initiated change in K-12, a report (*Action Plan*, 2006) from the Commission on the Future of Higher Education revealed that the American college system needed drastic change. The Commission categorized the change required as moving from ña system primarily based on reputation to one based on performanceò (*Action Plan*, 2006, para. 1). U.S. Secretary of Education Margaret Spellings identified student performance and student learning as the most critical point of the college ranking system.

To address this need, Secretary Spellings announced an action plan to make post-secondary education more ñaccessible, affordable, and accountable to students, parents, business leaders and taxpayersò (*Action Plan*, 2006, para. 2). The college level Action Plan echoed the NCLB mandates for K-12 levels. While colleges have typically been concerned about accessibility and affordability, the current emphasis on accountability is newer and remains to be implemented consistently in a more widespread way.

Huot (2007) found many inconsistencies with the findings of the Spellings Commission. He was concerned with the issue of accountability because it contradicts the greater role of faculty governance that is supposed to occur alongside it. Although Huot has taken issue with the vision put forth by Spellings, the legislation has been enacted, and postsecondary institutions have been struggling to interpret and comply with new initiatives.

One of Spellingsô ó Facts and Findingsô indicated that ñthe percentage of college graduates deemed proficient in prose literacy (able to read and extrapolate from a complex text) has declined from 40 to 31 percent in the past decadeò (*Action Plan*, 2006, para. 13). This finding has indicated that critical thinking/higher level/cognitive skills remain of paramount concern to the educational community. The goals of the Action Plan seem to come at critical time for colleges.

To initiate and implement the Action Plan goals, the U.S. Department of Education awarded grant monies in September 2007 for the purpose of determining ñreliable and valid measures for assessing undergraduate student learning across an array of learning outcomesò (*U.S. Department*, 2007, para. 1). Clearly, assessment of college students maintains a high priority.

From this federal mandate, regional agencies have begun to consider modifications to their accreditation processes. For example, the Southern Association of Colleges and Schools (SACS) passed significant changes to their process of accreditation by requiring institutions to assess ñthe extent to which it achieves these outcomesò (*Executive*, 2007, Section 3). The SACS changes specifically mentioned the area of student learning outcomes. One may infer that in the future, colleges that undergo accreditation will be required to address the assessment outcomes of their programs. The national and regional levels of post-secondary educational governance have seemed to indicate that colleges must move toward assessment of student learning as part of institutional accountability. Clearly, a movement toward evaluation and accountability has been initiated, and it will conclude at the department level and within individual classrooms.

Professional. Most undergraduate disciplines identify student mastery of writing skills and student mastery of cognitive skills as important goals, usually related to the specific content of the discipline (Appendix D; Cross, 1999; Lavelle & Zuercher, 2001; Ramey, VandeVusse, & Gosline, 2007; Sadler & Andrade, 2004; Westcott & Ramey, 1993). In college English, for example, undergraduate writing textbooks tend to teach writing in either prescriptive or descriptive ways, the descriptive way being the more process-oriented way of teaching it, and the prescriptive way being the more structure-oriented way of teaching it. Whatever combinations of goals or methods are required of college students by individual departments, professors evaluate improvement through student writing, and quite often, evaluation includes professor feedback and student revision of the essay.

In light of legislative mandates and accreditation requirements, colleges may increasingly need to demonstrate student achievement change or student mastery of outcomes. This means

that all professors and instructors may need to consider how to implement measurable change in individual courses and classrooms.

Taking into account federal legislation, regional mandates, and undergraduate goals, the University of South Florida (USF) developed the Cognitive Level and Quality Writing Assessment (CLAQWA) instrument (see Appendix B) in response to valid academic concerns about the writing skills of students (*Writing and Thinking Assessment*, 2007). This effort marked a decided change in the way campuses can handle the measurement of writing and thinking skills.

Until recently, the most commonly used rubrics for college writing have been holistic (Cooper & Odell, 1977; Davis, 1990), and their simpler rating scales of several points have limited their usefulness in measuring improvement. Other frequently employed writing scales involved grammatical analysis (Davis, 1989; Williams, 1990; Corbett & Connors, 1999), and their complicated analyses have limited their usefulness too, except to academic English. New directives require more flexible rubrics that can be implemented across a broader range of disciplines yet be as specific as needs dictate, so an assessment that measures both cognitive and writing skills may prove invaluable to many college departments.

Research-oriented. Several years ago, the University of South Florida (USF) initiated development of the CLAQWA instrument (see Appendix B) in response to their general education assessment that identified writing as a weak area on their campus (*Writing and Thinking Assessment*, 2007). Although they initially used a holistic scale, they developed the analytic scale in response to the limitations of the holistic scale. As trained scorers assessed student papers three times throughout the semester, they learned that the weakest student skills were cognitive. As a result, USF used test results to revise their general education curriculum so

that it would include process writing courses that emphasize critical thinking skills. The work of that campus coincided with the federal mandate that emphasized academic accountability, so the CLAQWA instrument may become a valuable tool to all institutions of higher learning.

Since the initial effort, the CLAQWA has been revised. Various versions have been developed to serve diverse purposes (Flateby, 2007). The CLAQWA can now be used by both students and professors in a variety of formats that serve numerous purposes. The instrument also holds value to institutions as they seek to identify, evaluate, and clarify curriculum goals.

The Cognitive Level and Quality Writing Assessment (CLAQWA) is a 16-point rubric, based on a 2-scale system, designed to evaluate both writing quality and cognitive levels. Professors can use the scale separately or in combination for writing assignments. The cognitive levels of the 2-point scale were derived from the work of Bloom (1984). The CLAQWA grouped the cognitive levels as follows: 1) knowledge, 2) comprehension, 3) application, and 4) analysis, synthesis, and evaluation. The writing quality assessment of the 2-point scale derived from commonly understood writing goalsð like unity, support, coherence, and sentence skillsð writing measurements frequently named in college writing textbooks. Both cognitive level and writing quality are evaluated on a 5-point continuum (Flateby & Metzger, n.d.).

The writing quality scale of the CLAQWA was developed by an interdisciplinary team and pilot testing. Faculty members were surveyed, and their needs were addressed. Consequently, the purpose of the assessment met needs from general education courses to specific courses in the major (Flateby, 2007). In addition, students were asked to write an essay explaining the features of the best course they had ever taken, and their needs were also addressed (Flateby & Eubank, 2008). A third purpose was to make the language of writing assessment clear to all potential users of the instrument. Although initially intended for use by

faculty only, the CLAQWA evolved into an instrument that could inclusively meet student, faculty, and institutional needs.

The University of South Florida drew its conclusions from several CLAQWA measurements: freshmen composition class reports, summary reports from exit classes, comparison reports of matched pairs of freshmen composition essays, individual student reports, and systematic peer reviews of writing (*Annual*, 2006). The researchers learned that freshmen writing cannot be compared to senior writing. Researchers at USF learned that writing across the curriculum reinforced writing skills learned in freshmen composition courses. The researchers learned that the CLAQWA instrument worked best when students can revise their essays (*Peer*, Oct. 2006, p. 1).

One aspect of college writing involved writing in disciplines other than English courses (*CLAQWA*, 2006). The CLAQWA enabled professors in all disciplines to give consistent feedback so that both students and professors could identify weak areas. Another aspect of college writing involved the cognitive level of the assignment. Every college writing assignment did not involve all of the higher order thinking skills; some college writing assignments only required lower order thinking skills. This instrument enabled professors to tailor the rubric to the assignment, helped them design the assignment carefully, and helped them communicate their cognitive expectations to the students (Flateby & Metzger, n.d.).

This researcher expected that student writing in the experimental group would show improvement when the CLAQWA instrument was implemented in the writing classroom. Implementation involved: teaching students about the cognitive scale and CLAQWA rubric prior to assignments, giving assignments that reflect such information, and evaluating student essays

according to the rubric throughout the semester. This researcher expected that student writing in the control group would not show as much improvement without the same intervention.

Research Questions/Problem Statements

Writing quality and cognitive performance have been inextricably linked (Flateby & Metzger, n.d., p. 4). The CLAQWA instrument was designed to enable instructors to make inferences about whether the use of this rubric influenced the cognitive level and the writing skills of college students. The goal of this dissertation was to determine whether students improved their writing skills and cognitive level skills as a result of the implementation of the CLAQWA rubric in a freshman writing course where the college expected that the culminating writing project of the semester should demonstrate such change. Before asking the research questions, inter rater reliability was ascertained from these two questions:

What is the difference between the scores from the two raters on the diagnostic essay?

What is the difference between the scores from the two raters on the final essay?

After determining the inter rater reliability, the following research questions were asked:

$RQ_{1:}$ What is the difference between the scores of the experimental group and the control group on the diagnostic essay?

$RQ_{2:}$ What effect does the use of the Cognitive Level and Quality Writing Assessment system (CLAQWA) instrument have on the writing skill of college students?

$RQ_{3:}$ What effect does the use of the Cognitive Level and Quality Writing Assessment system (CLAQWA) instrument have on the cognitive level of college students?

$RQ_{4:}$ What is the difference of scores between females and males on the diagnostic essay as graded by either rater?

$RQ_{5:}$ What is the difference of scores between females and males on the final essay as

graded by either rater?

In order to answer what initially seemed like a simple question, several hypotheses were investigated to arrive at a credible result. Inter rater reliability was determined before the comparison of final essay scores by looking at the significant difference between the scores from the two raters on the diagnostic essay and on the final essay. Gender differences were also addressed. The following null hypotheses were stated:

H_1: There is no significant difference between the scores of the experimental group and the control group on the diagnostic essay.

H_2: There is no significant difference of word skill scores between the experimental group and the control group on the final essay.

H_3: There is no significant difference of cognitive level scores between the experimental group and the control group on the final essay.

H_4: There is no significant difference of scores between females and males on the diagnostic essay as graded by either rater.

H_5: There is no significant difference of scores between females and males on the final essay as graded by either rater.

Research in composition studies has seemed to indicate that instructional guidance from the professor ð neither overt domination nor total lack of direction ð can enable students to improve their cognitive level and writing quality skills. The benefit of the CLAQWA instrument is that it standardizes, categorizes, and clarifies many writing outcomes that professors generally agree upon. Using the CLAQWA may give the professors who teach the experimental groups a common basis for communication about student outcomes, and as a result, see greater gains in their students.

This research assumed that college students, whose writing was evaluated by the CLAQWA instrument, would score higher on their final papers than college students, who were not taught by the same method of writing (See Appendix B).

Overview of the Methodology

This research study was a quasi-experimental, nonequivalent control group design. It was an evaluative type of research. The subjects were freshmen students at an urban, Midwestern state university. They were assigned to predetermined groups (Writing I class sections) through the college registration system. All of the students wrote at least three major essays, whose guidelines come from the English departmental syllabus. In addition, students were given numerous and varied types of shorter writing assignments, according to professor preference. They were expected to show improvement in the final essays as a result of completing the writing course.

In this research study, college freshmen from ten different sections of Writing I courses participated. Five courses, taught by the experimental professors, taught and evaluated students using the CLAQWA rubric. Five courses, taught by the control professors, neither taught nor evaluated students using the CLAQWA rubric.

The writing quality and cognitive level quality of the final papers of the two groups of students were compared. Two independent raters scored the final papers according to the CLAQWA rubric. The SPSS data analysis program was used to assess the scores. This research study investigated any statistically significant differences in the final paper scores. Although the first/diagnostic essays were also compared, their comparison served a different purposeð to determine the equivalence of the experimental group and the control group. The differences between scores on the final/third essays only compared differences in methodology.

Professional Significance of the Problem

The shift that has occurred in American education concerning academic accountability is considered by many to be monumental and far-reaching. It is the beginning of a new directional emphasis in education at all levels that has yet to be fully implemented. One obvious aspect of academic accountability has been the measurement of outcomes. Both the steadfast emphasis on critical thinking skills and the logical assumption that student writing reflects cognitive levels have contributed to the way educators perceive accountability. While many instruments have measured many aspects of all types of outcomes, a standardized instrument that reflects common understandings of writing skills and cognitive levels may become necessary to the implementation of what seems to be the coming trend in American undergraduate education.

The Cognitive Level and Quality Writing Assessment (CLAQWA) instrument was developed specifically with college level learning in mind. Although the CLAQWA was initially developed for faculty, its revisions quickly included institutional objectives and student functions. One may expect that just as federal education policy leads at the regional and state level, so will regional and state bodies lead departments and professors to similar ends. College professors, who have typically had little training in the area of education, may be expected to clarify their course level goals more articulately. The use of the CLAQWA may give all professors in all disciplines a common ground for talking about student outcomes. Besides, the CLAQWA has seemed to complement the methods and goals of freshman composition courses. By using the CLAQWA , this study hoped to indicate that first-year college writing could improve through the use of a specifically designed instrument and, as a result, enable departments and campuses to further implement institutional goals.

It is hoped that this research study will make a contribution to the implementation of changes within individual college writing classrooms. While many anecdotal studies report on diverse methods of success, none have yet addressed the specific requirements of accountability arising from recent legislation like the Action Plan for Higher Education (*Action Plan*, 2006) from the Commission on the Future of Higher Education.

Definitions of Key Terms

The following terms have been defined for the sake of clarity when referring to many types of practices within college writing classrooms.

CLAQWA: Cognitive Level and Quality Writing Assessment instrument; it is a 16-point rubric that measures two scales: cognitive skills and writing skills. Each point is evaluated on a 5-point continuum; the points on the rubric are selected to match the assignment

Cognitive/cognition: 1) refers to thinking skills, often as labeled critical thinking skills, or higher order skills; 2) refers to a progression from rote knowledge to applications of that knowledge; and 3) refers to a writing model that teaches writing with a cognitive emphasis

Descriptive writing: refers to a writing model that teaches writing with a lot of emphasis on process and less emphasis on models and structures

Prescriptive writing: refers to a writing model that teaches writing with a lot of emphasis on models, structures, and forms and less emphasis on process

Organization of the Dissertation

This dissertation about college composition has been organized according to the following pattern. First, this dissertation presents an introduction to the research design, its historical background, the research question and its professional significance, and an overview of

the methodology with its definitions. Second, this dissertation presents the review of college composition literature with its theoretical underpinnings and empirical examples, a summary of the direction of those studies, and the relationship of those studies to this research study. Third, this dissertation presents the methodology for the research. Included in this section are the type of research, the research context and site, the participants, and the instruments and materials. Fourth, this dissertation presents the findings of the research, with data charts, rubrics, permissions, and data analyses. Finally, this dissertation discusses the results of the study, the conclusions drawn from it, and recommendations for future research.

Chapter Two: Literature Review

A substantial amount of literature has been written about many aspects of college composition, and the search process for the material of this dissertation was conducted primarily through college databases like the Academic Search and Education databases. This dissertation review of literature about college composition has been organized according to the following pattern: theoretical and empirical. The review of literature presents an overview of the theory behind composition studies and categorizes the labeling of terminology in the field of composition studies for the sake of clarity. Included under this section, the review of theoretical literature presents: 1) theories about the cognitive process; 2) theories about post-adolescent cognition; and 3) basic principles for undergraduate education. Some anecdotal studies are also considered in this section. In addition, the review of literature presents empirical literature related to the study of college composition. These individual studies are classified according to the terminology previously mentioned and grouped according to similarity of ideas or practices. Finally, the review of literature offers a summation on the meaning of all this data and how two primary factors (information about teaching and legislation about educational measurement) may impact one another at the college level.

Review of Theoretical Literature

Accrediting bodies and campuses have mandated assessments of student learning; as a result, colleges are concerned about measuring academic quality and outcomes of students. A general assumption has existed that there is a close link between writing skills and cognitive skills, so colleges have seen a need to accurately measure both skills at the same time in

undergraduate student work. Because colleges have frequently required two semesters of undergraduate writing as general education core requirements, freshmen writing courses have constituted a logical place to evaluate such change in student writing.

In order to begin looking at composition studies, ideas previously investigated (like principles for good undergraduate teaching, numerous teaching methods, and thought processes of undergraduates) should be highlighted. Specific, measurable data has been analyzed (Hillocks, 1986). The Hillocks text has marked an important beginning to the enormous task of developing a bibliography and of synthesizing what has been learned from composition studies. It has also set in motion the establishment of the use of experimental and quasi-experimental designs in the area of academic English, a practice rarely seen in previous decades. More generalized practices have been collated (Roen et al., 2002). This book has represented encouragement, interpretation, and new directions for many pedagogical endeavors that occur within the freshman college composition classroom.

Additionally, the thought processes of college students and practical suggestions for undergraduates have been studied (Chickering & Gamson, 1987; Perry, 1999). Perry has used psychological theories to explain the behavior, motivation, and thinking of college students before he has indicated that professors should apply these aspects of intellectual development to individual teaching situations. Chickering and Gamson have very succinctly summarized seven principles for professor consideration that can be applied in student-faculty settings, including student/faculty contact, student reciprocity, active learning techniques, prompt feedback, focused time on task, high student expectations, and mutual respect (1987).

With the perspectives of accreditation, teaching, and learning in mind, Ramey et al. (2007) have articulated the idea that a systematic way of improving college student writing will

best serve campus, faculty, and student needs. Accordingly, when professors introduce a rubric to students at the beginning of the term, when professors give frequent feedback of writing to students, when professors explain to students how a rubric functions during the writing course, and when professors achieve consensus on essay standards (p. 70), the overall experience of the writing course could help students improve their writing (p. 71). This research project recognizes that the Cognitive Level and Quality Writing Assessment (CLAQWA) (Flateby & Metzger, 2001) is a rubric that was designed for use in the above-mentioned way, and this research represents an experimental design consistent with the research suggestions.

History and terminology. To gain a sense of direction about composition studies, the history of composition studies might be a logical place to start. Recently, Duncan (2007) briefly outlined the history of paragraph teaching/composition studies, and he covered material from the last two centuries. He asserted that although labels have changed, three predominant methods of teaching composition have continually prevailed in the teaching of composition courses. To separate and clarify many different labels, he categorized the teaching of writing according to three writing models that he labeled prescriptive, descriptive, and cognitive. (see the Definitions of Key Terms in Chapter 1). These three labels have been used throughout this dissertation in the defined ways, more completely explained in the following paragraphs.

To begin the discussion, prescriptive methods of teaching composition have depended on a concrete conception of the structure of writing. Topic sentences, paragraphs, and writing structure have played a major role in the prescriptive method. Alternatively, descriptive methods of teaching composition have deemphasized the structure of writing. As long as the writing has fulfilled the function of the writerôs intention, the writing has been considered adequate. More recently, cognitive methods and process models of teaching composition have emphasized

critical thinking skills, a psychological construct that is now beginning to be investigated more fully.

Furthermore, cognition has become an aspect that is now beginning to be considered as the more important aspect of writing. According to Duncan (2007), prescriptive and descriptive methods have tended to cycle up and down in popularity for the last two hundred years while cognitive methods have been a relative newcomer to the realm of composition theory. Even so, these three labels have remained helpful in sifting through material about composition studies.

Duncan would like to see all three methods coalesce and collaborate with other disciplines in a way that both unites old ways of thinking and enhances newer classroom practices (p. 487). Yet, he has correctly perceived the difficulty of teaching the more nebulous concepts that belong to the descriptive and cognitive aspects of writing; the popularity of prescriptive methodology has dominated due to its more tangible aspects. It may be that the educational paradigm shifts brought about by the NCLB Act and the Commission on the Future of Higher Education will force educators to accomplish this challenging and lofty goal in composition studies.

Much research in composition studies in the last several decades has forged a general perspective about the content of composition theory (Bloom, 1984; Chickering & Gamson, 1987; Duncan, 2007; Hillocks, 1986; Hillocks, 1995; Perry, 1999). The history of composition studies, research about composition studies, theories about college student learning, and applications of those theories within individual courses have covered the main areas of concern in composition studies. Most of them have indirectly inferred either the prescriptive, descriptive, or cognitive methods, and all of them have assumed that word skills measure cognitive levels. Clearly, a need still exists for a way to link many parameters in a more concrete and identifiable way

Cognition. The now, well-known measurements of cognitive growth have stemmed from the six levels of the cognitive taxonomy, as defined by Bloom (1984, pp. 201-7). He gave the most emphasis to the aspects of knowledge, the first level of cognition. The knowledge level included both the more concrete recital of facts as well as the more abstract recollection of universals, interrelations, or patterns (p. 62). ñThe knowledge category differs from the others in that remembering is the major psychological process involved here, while in the other categories the remembering is only one part of a much more complex process of relating, judging, and reorganizingò (p. 62). In other words, the levels of comprehension, application, analysis, synthesis, and evaluation may or may not be appropriate in a given educational situation; the needs of the situation may dictate which level(s) to use. It is commonly understood that every piece of writing does not involve all levels of the cognitive rubric, yet student comprehension of cognitive levels may help students understand different goals for different assignments, thus enhancing cognitive development of students. Also, student comprehension of cognitive levels may help to pinpoint areas where change is most needed. Bloomôs identification of cognitive levels has aided both students and professors.

To further explain the nature of cognition, Perry (1999) extended and developed the cognitive model for the undergraduate level of learning. He understood that cognitive development from concrete experience to abstract functioning repeated itself at older levels of development (p. 32). In other words, every time an adult begins to learn a new concept, the person (who may be mature in other ways) necessarily also falls back to a more concrete level of understanding before the person can move forward to more abstract functioning at the new concept.

In addition to the changeable, cyclical way that humans learn new concepts, other factors, not necessarily developmental factors, have influenced the way campuses think about cognition. For example, Perry observed that the movement toward diversity on campuses accompanied a movement toward relativity in knowledge, yet he viewed cognition as a growth process, not just a change process (pp. 2-5). This neutralizing tendency toward knowledge at the collegiate level has tended to diminish or overlook the importance of cognitive growth patterns. In other words, philosophical constructs can overtake psychological constructs, but returning to basic psychological concepts of growth may be a first step in reordering such omissions.

As a result, Perryôs efforts to pinpoint specific aspects of the cognitive process have helped college level educators rediscover and acknowledge the importance of the developmental steps in order to evaluate cognitive growth. His cognitive scheme has opened up ñthe possibility of assessing, in developmental terms, abstract structural aspects of knowing and valuing in intelligent late-adolescentsò (p. 16). To enable this evaluative process, he outlined four stages of undergraduate thinking.

Perry theorized that college students go through distinct steps of cognitive maturation which he labeled dualism, multiplicity, relativism, and commitment. To explain the steps, ñdualismò meant that students came to college thinking in a black/white, right/wrong, true/false, or a good/bad frame of mind. Initially, a college freshman believes that answers to complex problems should be as obvious as answers to math problems. ñMultiplicityò meant that students came to realize that many answers exist to solve a problem. They can understand that every authority is not necessarily competent, yet they still believe that an absolute answer exists somewhere.

ñRelativismò meant that students abandoned their faith in the perspective of dualism. ñCommitmentò meant that students accepted a particular position and learned how it affected their lives. Due to the introductory nature of all freshmen level courses, one would expect to see some cognitive growth in students, but probably not all four levels of it. The application of Perryôs thinking about college student cognition has been an appropriate beginning point for many composition studies.

Applications of cognition. In response to comments from higher education circles and to answer questions about how to help college students learn more effectively, Chickering & Gamson (1987) offered principles for undergraduate education based on fifty years of research in education. Information from state education agencies helped Chickering and Gamson distill much research into seven identifiable principles: 1) contact between students and faculty; 2) reciprocity and cooperation among students; 3) active learning techniques for students; 4) prompt feedback and assessment for students; 5) focused time on task for students; 6) high expectations from the professor; and 7) multiple ways of learning for students.

Although these principles have now been disseminated for more than two decades, many colleges still consider their advice current, perhaps because the authors ñaddress the teacherôs *how*, not the subject-matter *what*, of good practice in undergraduate educationò (p. 4). In other words, it is generally recognized that beginning this task of more effective undergraduate learning has meant starting with generalizations about many specifics.

Whereas Perry (1999) described the steps of cognition (dualism, multiplicity, relativism, commitment), Chickering and Gamson (1987) gave summative recommendations (student/faculty contact, student reciprocity, active learning techniques, prompt feedback, focused time on task, high student expectations, and mutual respect) that support the efforts of

campuses to move in the direction of teaching more cognitively. Colleges have still considered the advice of Chickering and Gamson to be sound today, even though some of the terminology has changed. To restate their advice for a college writing course, one might say that their advice has leaned toward a balance of prescriptive and descriptive methods, with a slight emphasis on cognition.

As the concepts of cognition have become more widely disseminated, researchers have shown interest in how to apply them in composition courses. In one cognitive study, DuBoulay (1999) investigated the relationship between reading and writing skills. In other words, she tried to bridge gaps between the steps of cognitive maturity through specific classroom strategies. She recognized the lack of critical reading skills among some higher education students and called attention to the fact that deficient reading skills have been closely linked with deficient writing skills.

DuBoulay offered three reasons why students face such academic difficulties. Her first reason, that the ñthe students themselves are the problemò (p. 148), reflected Perryôs cognitive theory that students have not yet progressed to the more abstract level of thinking required by college level material. Her second reason, that methods of ñteaching, assessing and evaluating performance are not relevant or appropriateò (p. 148) echoed many current arguments in the educational arena. Her third reason, that the formality of academic discourse is completely different than other types of reading familiar to students, returned back to Perryôs ideas about cognitive maturity.

DuBoulay suggested three strategies to help students improve critical thinking skills: read a text selectively rather than from beginning to end; annotate the text in a way that highlights statements of contrast or summary; and identify context clues that indicate the authorôs topic,

argument structure, and attitude. Overall, these strategies have helped students identify features common to academic discourse, and the recognition of such elements has preceded the cognitive growth desirable among college students.

In another cognitive study, Elder (2000) outlined nine strategies essential to the development of critical thinking, stages that students must go through in order to become mature thinkers. Her initial premise, that students cannot develop their critical thinking skills until they accept that a lack of them is indeed a problem that must be solved, has often been overlooked in modern education. Her secondary premise, that the development of critical thinking skills belongs to the skill set of everyday adult living, has coincided with Perryôs developmental ideas about post-adolescent maturation.

Her steps have helped students organize their time and manage their thoughts because the practical advice that professors can give to students may enable them to develop critical thinking skills. Elder & Paul (2001) explained more completely that critical thinking means thinking ñwith some set of ends in viewò (p. 40). In other words, when students were asked to restate the purposes of an assignment in their own words, this reiteration strategy helped them develop their critical thinking skills.

Elder & Paul (2002) focused on two elements of reasoning, specifically inferences and assumptions. In their view, when students began to identify and then question their inferences, they could better understand whether or not their assumptions were justified. As they asked questions to think through an issue, they developed critical thinking skills. These specific suggestions about cognition have held practical application for the college classroom, but they have dealt more with initial strategies than with final outcomes.

The studies mentioned above have shown how colleges have begun thinking and applying principles of cognition to the classroom setting. Initiating this effort has been tremendously valuable because most research looks back to preceding endeavors before building upon them. Without a doubt, cognitive theory has been clearly related to the teaching of freshman writing because writing is the primary means of ñassessing, encouraging, and grading student thoughtò (Flateby & Metzger, n.d.). The fluid, flexible natures of cognition and maturity have enhanced the potential of initiating intellectual growth in the college learning situation. Also, because freshmen writing courses have frequently been required at the beginning of the undergraduate experience, it has seemed logical to assume that the most accurate measurements of cognitive growth could be obtained at that juncture in student writing.

Background summary. In order to determine the efficacy of myriad writing treatments and to categorize teaching methods more cohesively, Hillocks (1986) completed a comprehensive overview of composition studies, compiled mostly from journal articles dated between 1963 and 1982. In this volume, he synthesized available research about writing treatments and categorized the teaching of writing according to writing models and focus of instruction. Whereas in the first volume, Hillocks (1986) conducted a meta-analysis of over 500 experimental treatments in writing classes, in the second volume, Hillocks (1995) offered methodologies for the teaching of writing based on the information of the first volume.

Of the 73 (out of the more than 500 journal articles about writing) that were suitable for statistical analysis (Hillocks, 1986, p. 187), he analyzed specific features to determine what worked best in teaching composition. He organized the studies according to four modes of instruction (writing models) and six foci of instruction (Hillocks, 1986, p. 192). Foci of instruction referred to ñthe study of traditional grammar, work with mechanics, the study of

model compositions to identify features of good writing, sentence combining, inquiry, and free writingò (Hillocks, 1986, p. 204). Hillocks defined modes of instruction as presentational, individualized, environmental, and natural process.

Of the modes of instruction, ñpresentationalò meant that the teacher dominated the classroom. ñIndividualizedò meant that the professor met frequently with students in individual conferences and engaged in traditional classroom instruction very little. ñEnvironmentalò indicated a balanced classroom setting. Students, teaching materials, teacher roles, and learning tasks functioned in a balanced way (Hillocks, 1995, p. 221). For example, students may have lead small group discussions in which they focused on solving specific problems according to guidelines from the instructor. ñNatural processò indicated a student-centered classroom setting. For example, students may have written on topics of their choice. They received feedback from peers, but the instructor did not present rules, criteria, or models of writing. Students revised as they wished, and student-led discussions avoided structured problem solving. Hillocks (1986, p. 199) rated the four teacher roles from best to worst: environmental, natural process, individualized, and presentational. Although the effects of the environmental mode were significantly greater than the other three modes, the differences between the other three modes are not significant. In other words, students have shown improvement when taught in any mode.

Hillocksôs presentational mode of instruction seems to correspond roughly with the prescriptive model of Duncan. Hillocksôs natural and individualized modes of instruction appears to fall generally under the descriptive model of Duncan. Hillockôs environmental mode of instruction seems to include the descriptive and cognitive modes of Duncan. ñEnvironmentalò also appears to reflect Chickering and Gamsonôs idea of the balanced classroom and seems to take into account some aspects of cognition.

Duncan appeared to disagree with Hillocks that no particular approach had been shown to be more effective than another (Duncan, p. 480). Duncan advised that when all writing methods intersected in a positive way in the freshmen writing classroom, students demonstrated greater writing skills than when they were taught according to one method only. A closer look at the definition of ñenvironmentalò has revealed that it really means that the teaching situation balances ñstudent, materials, activities, teacher, and learning taskò (Hillocks, 1995, p. 221) within the classroom. So, Duncan actually seems to agree with Hillocks that varying the dimensions of instruction improved student skills more than a single approach to instruction. Both have leaned toward the recommendation that a combination of approachesð prescriptive, descriptive, and cognitiveð provided the best classroom situation for learning.

The thinking of both Hillocks and Duncanð one who has studied writing treatments and one who has assessed historical trendsð have seemed to say that when professors gave a framework to the writing assignment, the quality of student writing may have shown improvement. Professor-led involvementð neither domination nor absenceð in the learning process is critical. Their thinking has reflected age-related cognitive theory about college students explained by Perryð that appropriate classroom management can foster cognitive growth revealed through student writing.

When taking into account several factors: 1) legislative and accreditation directives (*Executive*, 2007; *Four Pillars*, 2004); 2) the history of composition studies (Duncan, 2007); 3) theories of cognition and their application the college level (Bloom, 1984; Perry, 1999); 4) summaries of writing practices (Hillocks, 1986); and 5) summations of advice for undergraduates (Chickering & Gamson, 1987), one begins to see a need for a method or a

program or a tool that may incorporate all aspects of a complex situation into a cohesive unit and allow valid measurements of its individual components.

With new demands and requirements in mind, Flateby & Metzger (2001) from the University of South Florida developed the Cognitive Level and Quality Writing Assessment (CLAQWA) System, an assessment tool that measured both ñwriting skills and cognitive levelò (p. 4). The CLAQWA rubric was developed at the University of South Florida as an instrument whose intention was ñto help instructors standardize their evaluation of writing and assess the cognitive level attained in student writingò (Flateby & Metzger, n.d, p. 2). The rubric has been in development since 1999 (Flateby & Metzger, 2001, p. 4), and it helps both students and professors clarify writing objectives. The instrument was developed by an interdisciplinary team and pilot testing. The CLAQWA has undergone refinement and revision as a result of surveying faculty and students (Flateby, 2007). Because this rubric defines and standardizes the skills involved in the writing processes and outcomes, it allows a valid comparison between two sets of final papers.

It is an instrument that has fulfilled a genuine need. The CLAQWA has defined writing skills and cognitive levels, giving professors a rubric for evaluation. To use the CLAQWA correctly, professors have presented an explanation of cognitive levels, assignment expectations, and requisite skills to students before an assignment is completed. When students have gained an understanding of the level of work required for the assignment, it may be assumed that their work has reflected this increased understanding, i.e. showed improvement.

All in all, the perspectives of legislation, psychology, history, and pedagogy seem to have drawn the same conclusionð that careful planning may produce the greatest skill improvement in the writing of college freshmen. If student writing holds the potential to indicate cognitive

growth, then an assessment tool that measures both cognitive level and writing skills may prove invaluable to the undergraduate situation by enabling educators to begin moving in the new direction of accountability.

Review of Empirical Literature

The research in composition studies has reflected a wide array of approaches, indicating the many areas of concern in writing as well as the perplexities of measuring both cognitive level and word skills. In addition, academic English has generally used a different vocabulary than the discipline of education, so the recent passage of legislation has made it more important to simplify the terminology in order to recognize patterns. (see Definitions of Terminology in Chapter 1). Most composition studies have assumed that word skill change indicates writing change (i.e. cognitive growth), but they have not necessarily stated it as such.

As previously mentioned, writing quality and cognitive levels have always been closely connected. Many composition studies have employed a rubric or a program or a method and have claimed that it improved the quality of writing among college students (Corbett & Connors, 1999; Davis, 1989; Friend, 2001; Hafer, 2001; Oldenburg, 2006; Roen, 1984; Sanders, 2001; Sanders & Littlefield, 1975; Shuman, 1991; Soles, 2005; Williams, 1990). Additionally, writing models, classroom techniques, instructional strategies, classroom suggestions, personal practices, and even technology have been implemented in composition studies (Cazort, 1982; Davidson et al, 2002; Dwyer, 1992; Eades, 2002; Ford, 2002; Horning, 1997; Roen et al., 2002; Shelley, 1998). These studies have revealed a past, present, and continuing concern with similar writing issues. Now, federal, regional, and local mandates have made their further refinement and clarification more readily apparent.

Grammar rubrics. The most widely used writing assessments have involved grammar because grammar has remained such an obvious component of writing, and one facet of grammar has involved the use of transitional words and phrases for coherence, i.e. better comprehension of material. Jacobs (1977) considered the analysis of coherence to be so vital that she labeled it, ñgrammar beyond the sentence levelò (p. 10).

For example, Roen (1984) reworded paragraphs of published articles and gave them to students before testing them for comprehension. Some paragraphs were rewritten without any transitions, some paragraphs were rewritten with a lot of transitions, and some paragraphs were rewritten with an average number of transitions. He learned that too many transitions overloaded the passage to the detriment of student understanding while the other two rewritten passages allowed similar and better recall of material.

As a result, he advised professors to teach students to use transitions, to connect words/phrases, and to use conjunctive adverbs carefully. Few would disagree that college handbooks make coherence seem like a rote rule method with their lists of transitions, but in reality, ñgood writing requires good thinkingò (p. 36). In other words, coherence involves using words thoughtfully and purposefully. Nevertheless, coherence is somewhat difficult to measure apart from rote lists. Although the study of Roen has highlighted one aspect of writing, most professors have seen a need to measure student writing more comprehensively.

In another coherence study, Jacobs (1977) spent a lot of time taping student conference conversations for the purpose of helping students develop coherence in the final drafts of their assignments. She found that when conversations with students digressed away from grammar toward subject matter, the students began to see what they had omitted to write. Focused

conversations clearly helped students improve coherence, not as a mechanical exercise, but as a thoughtful process. She offered a schema of coherence for classroom use.

Jacobs referred to models of coherence, as did Marzano (1983), but Marzano identified cohesion separately from coherence. He saw coherence as a set of rules applied to the whole essay, while cohesion gave a sense of unity or flow to the essay, perhaps apart from the use of familiar coherence devices. Finally, Marzano suggested quantitative techniques for coherence analysis.

Marzanoôs understanding of cohesion and coherence has seemed to lean toward a more descriptive form of writing composition whereas Jacobsôs understanding of coherence has seemed to lean toward a more prescriptive form of writing composition. Even when researchers have tackled the same quality of student compositions, they may differ significantly in their approach to the topic. Nevertheless, the door has been opened up for future studies to determine growth or make comparisons of student writing, based on identifiable components of writing.

In addition, Davis (1989) indicated that teaching the sentence-combining technique of Christensen contributed to improvement in overall quality of freshman writing. According to Davis, Christensenôs rhetorical technique involved the use of free modifiers, descriptive phrases/subordinate clauses that are very detailed modifiers placed appropriately in a cumulative sentence, to supplement meaning and add professionalism to the essay.

Thinking in a similar vein, Williams (1990) made several suggestions to improve written communication skills: adhering to the standard American English sentence order of subject-verb-object; keeping the topic consistent; using the fewest words possible; and employing subordination, coordination, and parallel structures for readability. He aimed to identify clear writing by both cohesion, i.e. ñflowingò words and coherence, i.e. ñwell-formedò words (p. 101).

He argued that the differences between ordinary speech and formal discourse have stemmed from historical and cultural influences (p. 3) that have unquestionably existed as part of the English language heritage; this stylistic flexibility of the English language has created problems when teaching writing. Williams (1997) summarized his suggestions into ten keys for clear writing and offered practice exercises for students. This later volume was conceived as a workbook for students whereas the earlier volume was written as an explanation of principles for practitioners.

Similarly, Shuman (1991) understood that while students who learn a lot of grammar do not necessarily produce a significantly better quality of writing, he agreed with theorists who proposed a middle ground kind of way to teach grammar instruction. Because students tend to repeat a small number of similar errors when they write, he has considered it better to ñestablish grammatical prioritiesò (p. 82) that accommodate those needs. In other words, when the professor focuses on limited grammatical correction issues, students will be better able to understand how to improve the quality of their writing.

In like manner, Corbett and Connors (1999) emphasized types of sentences, sentence openers, and diction as ways to improve writing. They have explained that the teaching of classical rhetoric entailed the teaching of a sequence of assignments with increasing difficulty, and they have suggested that modern teaching reflect a similar sequence even though modern education has scaled down the schedule of teaching such courses. Corbett and Connors asked students to analyze a published paragraph according to a detailed stylistic chart (p. 370) and then compare that writing with their own. They found that when students marked both pieces of writing, they were surprised to learn how differently their own writing compared to the other

piece. This revelation helped them to change the way they were writing, and as a result, they improved their essays.

Likewise, Soles (2005) identified the stylistic features of research papers from exemplary first-year college writers. He employed the style charts of Corbett to assess word choice, verb usage, voice, subordination, and vocabulary. According to Soles, the best first-year writers observed the conventions of standard American English; they used appropriate vocabulary, they followed the subject-verb-object sentence order, and they employed various beginning strategies to add interest to the writing. His work simply recorded the features of better quality writing. According to the above-mentioned writers, exemplary freshmen writing exhibits specific, identifiable word skills, and all of these writers view word skill change as measurable.

Grammar and structure. In light of the recent emphasis on measuring outcomes, Sanders (2001) suggested using alternative assessment strategies, in the belief that alternative assessments are focused more on learning outcomes than on teaching methods. She recommended generating assessments that were more creative than the standard, familiar types of assessments. Instead of professors who simply teach through lecture and test on the material covered, professors should think in a more creative way to get students to employ critical thinking skills about the same material. Students have always needed the basic concepts of a subject, but basic concept knowledge should lead into higher levels of information processing, and this is where alternative assessment strategies have helped students improve cognitive skills.

According to Sanders, professors can replace standardized tests, final exams, short quizzes, and course essays with performance tests, journals, portfolios, summative assignments, cooperative activity groups, or alternative ways of grading. Not all types of assessments work equally well with every type of class, but as long as assessments are clearly linked to Bloomôs

taxonomy, Sanders has seen no valid reason for professors to continue to use standardized assessments unthinkingly. Generally speaking, she has recommended that assessments be developed thoughtfully, so that their purposes are clearly fulfilled.

As indicated by Sanders, orienting the assessments to the students has shown the ability to increase their effectiveness, a concern to the mandates for the NCLB generation. However, a huge gap exists between standardized testing procedures and classroom testing procedures as much as a gap exists between individual classroom practice and corporate group practice. When a researcher like Sanders (2001) suggested using alternative assessment strategies, it seemed like a good idea. But the question remains whether using an alternative assessment is just a better idea or simply a different idea. Educators have always considered assessments useful, but their efficacy has sometimes been questioned.

In spite of the current unpopularity of teaching grammar, grammar has not yet completely disappeared from the freshman writing curriculum. Oldenburg (2006) has still found it necessary to teach some grammar to college students. Rather than teach the lessons himself, he compiled a list of twelve common grammatical problems and asked pairs of students to study, explain, and teach the lessons to the rest of the class. Students used the grammar handbook to research the lesson and prepare a handout with examples and exercises for the rest of the class. The peer taught lessons were presented over a six week span, and as time went on, Oldenburg noticed that students learned to ñlocate the óclass expertô on a particular grammar point and ask for clarificationò (p. 75). From his perspective, students need to practice correct grammatical constructions, or they may continue to make basic errors in mechanics, errors that prevent them from presenting themselves in an articulate way to the working world, particularly when they go job hunting.

These several grammar and structure oriented studies have revealed a persistent concern with form more than with content, i.e. they have advocated a more prescriptive style of teaching college writing. And, it is undeniable that form has always been more easily measured than content. Nevertheless, structural concerns have remained unpopular as a primary measurement of writing; this perspective seems to reside in a modern outlook which holds a belief that an overemphasis on rules stifles quality.

Analysis of writing. Other commonly used, rubric-oriented assessments have involved two timed essays (Hillocks, 1986, pp. 156-160). The first essay was usually assigned before instruction began, and the last essay was usually assigned after instruction ceased. Both sets of student essays were graded holistically. Holistic scales are fairly simple, with a range of several numbers that indicate excellent to inadequate values. Validity was derived from the replication of the same learning conditions in each testing situation. Holistic scorers scanned rapidly to gain a sense of quality from the writing. While this pretest/posttest scheme has seemed adequate to some, others have questioned its validity (Hillocks, 1986, p. 155).

Accordingly, Sanders and Littlefield (1975) analyzed essay data from two writing courses, written during one semester. In the experimental group, the students wrote a researched freedom-of-choice pretest and a researched freedom-of-choice posttest. In the control group, the students wrote an impromptu pretest and an impromptu posttest.

Most prior research in composition had been done in the control group format, so in response to concerns that randomly conducted pretest essays and randomly conducted posttest essays showed no significant differences in overall writing quality (Sanders & Littlefield, 1975, p. 145), the researchers desired to learn whether significant improvement could be obtained

when students were given the same (research) conditions to write both the posttest as they were given to write the pretest and other essays/tests throughout the semester.

Sanders and Littlefield (1975) indicated that the validity of the impromptu, timed essay seemed questionable for several reasons. First, students were given no alternative about the choice of topic. Second, the inflexibility of the testing situation meant that students had no opportunity to revise written work, as they had probably been taught to do in the writing class. Third, the rigidity of the testing situation meant that slower writers could not perform as well as speedier writers due to the time constraint (p. 147). If the overall aim of the writing course emphasized a process style of the composition process, then the timed entry tasks and timed exit tests were actually asking students to write according to different parameters than those they had spent the entire semester learning.

Overall, the chief criticism of the holistic scoring method has seemed to center around the differences between testing conditions and teaching conditions (Sanders & Littlefield, 1975, p. 147), so the unpopularity of structural concerns as a measurement of writing quality may not be entirely valid. Even so, new assessment mandates have been causing educators to re-think what they are assessing as much as why they are assessing it. Holistic measurements may no longer seem adequate enough to fulfill recent mandates.

Many professors have used various types of models, techniques, or strategies to teach writing, not just the rhetorical models featured in many writing textbooks. All of these practices have contributed insight to the general discussion about college writing. At the same time, one cannot help but notice that they have been so specific that they have not necessarily seemed to belong to the same cohesive body of information. Clearly, a standardized way of talking about college writing may promote the pedagogy of the discipline more than any other factor.

Models/strategies/methods. To give one example of a model, Dwyer (1992) used journalism as a prescriptive model to teach college students how to write about the readings in class. His method asked students to use the inverted pyramid form of most newspaper articles. Students presented the most important information first, and then amplified it with adequate details and various sources. Since many writing courses have asked students to respond to readings in class as a type of writing practice for more formal essay writing, student s found this approach supportive of the larger effort of the course.

To give another example of a model, Horning (1997) believed that when professors taught reflective writing via portfolio writing, the quality of student writing would show improvement. Throughout the first semester of college writing, her students wrote ten papers of choice from a total of thirty prompts. The papers were brief, but the brevity of the papers boosted the self-confidence of students in revising and editing their essays. According to Horning, reflective writing, a descriptive model, has allowed students to consider many ways of revision and has given students insight into their own thought processes. Improvement depended upon the professorôs perception of student improvement.

From a different perspective altogether, Cross (1999) discussed the impact of the flood of information and research about college student learning. Instead of supplying a specific model to follow, she first explained that current times seem to be becoming somewhat dependent on what authorities advocate as best. She contrasted conventional teaching that is ñbased on a hierarchical model in which those who know teach those who do not knowò (p. 259), to many contrasting understandings, which maintain ñthat knowledge is constructed by humans through social interactionò (p. 259). Even though these two ways of thinking have dominated the thinking of society, Cross indicated that the twenty-first century should move beyond ñthe

authoritarian search for right answers or the egalitarian notion that all ideas are equally validò (p. 265). She has indicated that either of these perspectives has classified students into groups, which she has understood as defeating the purpose of education, which she has identified as more individual in nature.

Cross (1999) preferred to think that research initiated rather than answered questions. She advocated collaborative learning in which students develop their own answers rather than relying on getting answers from professors. She assimilated principles from Bloom (1984), Perry (1999), and Chickering and Gamson (1987) to say that college learning should be relative and contextual. To interpret her philosophical stance for the writing classroom, she has leaned toward descriptive teaching, while at the same time, she has left the specific application of it up to those who teach.

These three aforementioned models (journalism style, portfolio keeping, collaborative learning) have all contained merit for undergraduate courses, yet for the purposes of assessment, difficulties arise. While all of the above models have accomplished specific classroom purposes, a problem still lies in how to give an equable rating under such diverse teaching contexts.

The holistic commentary of Sanders and Littlefield has stood in contrast to many narrower studies, like Friend (2001), who singled out a solitary skill like teaching generalizations as a means to improve summary writing among low-skill college students. Her method has seemed similar to the journalism method of Dwyer, who worked with traditional college students. Friend labeled the terms differently, but clearly, professors have shown concern for similar issues facing students.

Like Dwyer, Friend linked the college writing course to the purposes of other college classes, a well-recognized connection. She applied a text-processing theory that advocates

argument repetition as the critical component in the cognitive process (p. 19). Her study, using a prescriptive method, indicated that direct instruction for this single skill helped students learn how to construct their own main ideas from a text, rather than looking in the text for the summary statement.

Enlarging on Hillocksô (1986) idea that instructor feedback is critical to student success in writing, Hafer (2001) linked a purchased system to success in freshmen composition classes. The system involved group study and peer-facilitated learning. Its descriptive goals complemented the goals of freshmen writing because the system emphasized the acquisition of writing skills. Structured group talking and collaborative group work were the most common practices of this program, and both of these activities aided students when editing and revising essays. The result of employing this system was greater peer contact and improved peer responses. In other words, writing skills of college students improved as a result of using the methods of this purchased writing system in class.

Thinking more about student feedback than instructor feedback, Eades (2002) used peer revision workshops in freshmen writing classes. The construction of her system involved ñstudent and instructor participation from three anglesò (p. 61). She referred to it as pedagogical triangulation; it addressed weaknesses of many peer revision systems. The students interacted with the instructor, with their peers, and with student panels to revise and edit their papers. She advised professors to think through the process very carefully before implementing it throughout the semester. According to Eades, ñpeer revision may best be achieved through a holistic approach to collaborationò (p. 65), but it needs to be well defined in order for students to benefit from it most.

With another specific practice in mind, Ford (2002) briefly related the effectiveness of teaching anecdotally in undergraduate classes. His advice, which is somewhat descriptive in nature, has reminded college professors of a forgotten tool of academia. He has maintained that students remember the point of an anecdote far longer when the point is made through a story, even to remembering the point of the anecdote many years after college is over (p. 3).

These five above-mentioned strategies (teaching generalizations, applying text processing theory, using purchased systems, practicing peer revision, teaching anecdotally) have offered diverse advice for the college composition classroom, so an obvious question remains concerning how to value one strategy over another. If professors can relate their goals to a standardized reference point, then the groundwork for discussion can be laid.

Two other studies (Davis, 1990; Mohammed & Jaber, 2008)) have taken a more cognitive approach and have highlighted differences in student learning outcomes; both offered less prescriptive results than previously mentioned studies. In a comparative study, Davis (1990) used two types of measurement to compare writing outcomes of form-centered instruction with outcomes of process-centered instruction (p. 3). Bambergôs Holistic Coherence Scale measured the former, and the Discourse Matrix measured the latter. In other words, he compared outcomes of prescriptive writing instruction with outcomes of descriptive writing instruction, but by two different methods.

The emphasis of form-centered instruction was writing according to a predetermined structure where models, outlines, and grammar held primary importance. The Holistic Coherence scale evaluated aspects of both cognitive level and word skill but relied on a simple 4-point scale for measurement. The emphasis of process-centered instruction was multiple revisions of an essay where peer feedback, teacher conferencing, and work sharing held primary

importance. The Discourse Matrix diagrammed the sentences students had written, emphasizing word skills.

According to Davis (1990), the Coherence Scale indicated that form-centered writing showed greater gains while the Discourse Matrix indicated that process-centered writing showed greater gains. This study has differed from other studies in that it evaluated student writing in more than one way. It suggested that different modes of instructionð either prescriptive or descriptive or cognitive methodsð in a writing course may initiate different types of changes in the writing of college students.

In another comparative study, Mohammed and Jaber (2008) gave insight into college writing from the English as a Foreign Language (EFL) perspective. They studied the effect of two writing approaches to teaching a specific grammatical point in EFL classes comprised of freshmen and juniors. They compared pretest and posttest results for two groups. The first group was taught deductively, by learning grammatical rules. The second group was taught inductively, by focusing on language context. According to the terms of this study, their deductive method would correspond to a prescriptive method, and their inductive method would resemble a descriptive method.

Mohammed and Jaber (2008) concluded that the prescriptive group showed greater gains than the descriptive group, but they attributed the reasons for the improvement to the fact that mature, highly motivated adult students have greater logical understanding than younger, less skilled students. Mature students have learned how to apply simple rules into other situations. Their results differed significantly from the recommendations of Cazort (1982), who summarized best and worst composition practices.

Mohammed and Jaber explained discrepancies in scores by identifying maturity as a predisposing factor to writing improvement. Their perception of students has seemed different than the perception of most American colleges, who may expect cognitive maturity to occur during the college years but do not necessarily expect to meet it in entry level classes.

Technology. Technology has also impacted composition instruction (Dave & Russell, 2010; Davidson-Shivers, Nowlin, & Lanouette, 2002; Shelley, 1998). From the introduction of computers into society in the 1980s to their more widespread usage by the 1990s, the advent of technology made professors consider how to incorporate computer literacy into the college writing classroom.

By the end of the last decade of the last century, Shelley (1998) underscored the importance of this formerly new skill. The last ten years have seen college students become more proficient technologically. Her advice, that first-year composition classes be taught in the computer laboratory to give students the best preparation for the rest of their college experience, has been implemented by many college campuses.

Although technology has been implemented in writing courses, Dave & Russell (2010) wondered whether the use of it has changed the way college students work at the drafting and revision aspects of writing. Their surveys suggested that while technology has made writing easier, it has brought about little change in student practice (p. 427). They have advised that writing processes and its relationships to technology be explored further.

While technology can never be considered a teaching method, it has nevertheless become a vital component of the methodology being used in the current composition classroom. Moreover, technology has further advanced enough to make all kinds of measurements easier to calculate. Word-based programs and number-based programs have lightened many workloads

that were formerly more cumbersome to professors, and perhaps they may aid in further research studies.

In the area of college writing for example, Davidson-Shivers, Nowlin, and Lanouette (2002) explained how they introduced multi-media lessons with a patented system to investigate whether computer-generated lessons improved the writing of college students (p. 20). They also administered Kolbôs learning style inventory to see whether learning styles held any correlation to writing performance. Their multi-media lessons taught the prewriting skills of brainstorming and outlining techniques, yet students had options: to move slowly or quickly through the lessons; to complete all or skip some of the extra practices for a lesson, if desired.

Student writing skill was measured by the final composition only; they were scored according to a 6-point holistic scale developed by their English department. The scale seems to incorporate cognitive levels and word skills, but in a simplified way. Researchers found that while students enjoyed the multi-media aspect of the course, students who wrote on a daily basis improved their writing skills more than those who wrote only when required. In other words, technology has affected neither learning style nor writing style as much as student effort. Students who completed all of the assigned work and who wrote regularly outside of class showed the greatest improvement in writing.

While technology is neither a model nor a strategy, it has affected composition instruction (Dave & Russell, 2010; Davidson-Shivers, Nowlin, & Lanouette, 2002; Shelley, 1998). The versatility of technology has increased its usefulness to writing pedagogy. In the future, one may expect to find even more varied technological applications for college writing needs.

Larger composition studies. All in all, many efforts have been made to summarize the extent of college composition studies, Hillocks (1986) being the most notable, analytical, and complete. With the intention of giving practical advice to less experienced college instructors, Cazort (1982) summarized the research about writing improvement methods and identified eight as being in widespread usage. Of the following: ñtraditional grammar, structural linguistics, transformational grammar, sentence-combining practice, frequency of writing, intensive correction, increased reading and precomposition experiencesò (p. 1), Cazort offered sentence-combining practice, increased reading, and precomposition experiences as most beneficial to writing improvement. He also asserted that traditional grammar benefits student writing the least. His advice leans toward more prescriptive methods of writing pedagogy.

In a broader collection, Roen & Pantoja (2002) compiled 93 essays on freshmen composition topics, including strategies. Their compilation has exemplified the diversity of writing strategies. For example, Rhodes advised avoiding grammar entirely and shifting the emphasis to any other facet of writing (p. 523). Hodges advised class discussions of literary texts to help students learn about writing quality (p. 534). Karolides explained how he used Christensenôs sentence-combining technique to help students connect their thoughts with their writing (p. 536). Similarly, Golson concentrated on parallel structure lessons (p. 551). Vaught-Alexander gave students activities that linked their writing to her lesson plans (p. 546). Licklider emphasized the importance of giving feedback to students and explained her particular way of doing it (p. 560). Kearns suggested a minimal way to mark papers so that students would take responsibility for their own errors (p. 567). One can see from this brief synopsis about writing strategies that effective classroom practices have included prescriptive methods, descriptive methods, and cognitive methods.

Departmental concerns. To underscore the problems involved in creating a college level freshmen English program, Westcott & Ramey (1993) highlighted the plight of many departmental decision-making situations. When the English department decided to change its freshman composition program, the department decided to teach it composition courses in four distinct sequences. The sequences were carefully designed to measure student progress. While the campus has been satisfied with the initial results of the changes, the department would still like to see further improvements such as: introducing a rubric early in the semester, using a rubric efficiently, communicating to students the importance of clear writing, increasing opportunities for student feedback, and achieving consistency of outcomes among faculty (Westcott & Ramey, 1993, p. 70). Perhaps a flexible yet unified framework, like the CLAQWA, could help colleges in similar situations clarify goals for students.

To focus on another side of departmental issues, Sadler and Andrade (2004) described how instructional rubrics could teach students the disciplinary skills necessary to success in college writing. They have compared the writing process to navigation and have explained a major difference between students who love to write and students who hate to write. Students who love to write were seen as experienced sailors, with hands firmly on the tiller to steer through rough waters. Students who hate to write were seen as inexperienced sailors, who have yet to recognize the purpose of a tiller or the need for steering correctly. When professors employed both self-assessment and peer assessment rubrics as the main correctional mode of the course, students achieved increasing proficiency in the writing process.

Likewise, Ramey, VandeVusse, & Gosline (2007) explained how the use of a departmental writing rubric both clarified and improved intended student outcomes in college writing. The specific needs of the Nursing department made their rubric more appropriate for

their own purposes than other standardized rubrics (p. 71). Their research has implied that writing rubrics serve specific yet limited purposes. Without a doubt, rubrics are seen as valuable tools to professors and researchers.

The information overload facing departments and professors has sometimes appeared burdensome. In response to an inundation of information that has advised college educators to ñmake student learning their top priority,ò Cross (1999, p. 255) questioned the usefulness of the information overload. She lightheartedly pointed out the difficulty of navigating the seven principles for undergraduates, the three critical conditions for excellence, the nine strategies for improving, and the twelve attributes of good practice (Cross, 1999, p. 256). Her jocular point has been well-taken, that in spite of the assortment or the difficulties of the choices, changes will nevertheless come, and professors must consider how students may learn more effectively in order to fulfill the new criteria.

Instead of looking at rubrics or measurements or goals from the departmental perspective, Lavelle and Zuercher (2001) worked from a psychometric perspective to gain an understanding of how college students approached academic writing tasks. They administered a Likert scale inventory about writing processes and followed up with individual student interviews. They learned that while the more articulate writers saw themselves as involved in the writing process, the more basic writers detached themselves from it by performing assigned tasks in a perfunctory way. It has seemed that student ñawareness of writingò (p. 384) contributed most to student motivation and subsequent success. Lavelle and Zuercher believed that professors could facilitate writing efforts by establishing a specific climate within the classroom. Setting cognitive goals, giving structural guidelines, emphasizing revision, giving appropriate feedback, allowing students to share efforts with the class may all facilitate studentsô efforts (p. 385).

In agreement with the principles of Lavelle and Zuercher, Dave and Russell (2010) advocated more research into student understanding of the writing process. Their research indicated that while technology has impacted the college writing classroom, student effort and perceptions about writing have not changed significantly.

Both of these perspectives about students have helped those who teach keep the nature of students of primary concern but may be more difficult to measure quantitatively. Clearly, departmental decisions, campus initiatives, and student processes all need to work together to accomplish the same end.

All in all, composition studies have employed various features of prescriptive methods, descriptive methods, or cognitive methods in writing instruction as a means of assessing change in college student writing. These studies and dozens of others have seemed to indicate that teaching ñspecifiable proceduresò (Hillocks, 1995, p. 223) may initiate the greatest change in student writing. Composition instructors, departments, and colleges are faced with a dizzying array of models, procedures, suggestions, methods, programs, theories, and ideas. It may well be that the introduction of a simpler, more unified framework will make it possible to navigate this river of information.

For the purposes of this research, such outcomes come from the English Department General Syllabus for the Writing I course at Youngstown State University (see Appendix D). As accrediting bodies and campuses mandate assessment of student learning, the emphasis on cognitive skills seems to override issues related to the efficacy of one method over another, yet cognition will be measured by writing.

Summary of what Previous Research seems to Mean

As the above-mentioned literature review indicates, college teaching is not as simple as it may initially appear. Before professors even enter the classroom to teach, much preparation and thought has already occurred. While professors may be more individually concerned with presenting content to students, campuses may be more broadly concerned with the external dynamics that affect measurable teaching outcomes. In future, it may be that links between the external dynamics and individual concerns will need to be made clearer. This study investigated whether the implementation of a specifically designed rubric could be a useful part of assessment, by bridging the broader requirements of the college and the narrower concerns of the individual classroom. The benefit of using this specific rubric is that it includes both cognitive level and writing skill measurements, a different enough approach that it may help campuses move toward more useful and accurate measurement of student outcomes.

How Previous Research Relates to this Study

It is well-known that trends in education enter the individual classroom through a lengthy process that originates at the federal level, then filters down through regional and state governance procedures, and finally arrives on campus to enter the departmental decision-making processes. In spite of the fact that trends can give rise to negative experimental tactics in classroom practice as much as they can bring about positive change in the classroom, when trends become mandates, educators struggle to meet the new criteria.

In addition to the legislative aspects of education, historical trends and empirical studies have played a significant role in the way that trends in education are perceived. Perhaps as the academic community pursues dialogue about instructional assessment, it will gather together a common consensus about how to implement it.

Nevertheless, there remains a need for an instrument that is flexible enough and comprehensive enough to be useful in many disciplines. Because the CLAQWA instrument evaluates both cognitive level and word skill in writing and also focuses on outcomes, it meets current needs. The 16-point rubric categorizes and standardizes the foci of instruction, another valid concern arising from mandates. As a widespread measurement, the CLAQWA instrument may prove to be an invaluable asset to instructors and campuses as they seek to define and clarify assessment goals.

The Conceptual Framework

The conceptual framework for this research proposal involves two designs: the overall scheme and the narrower scheme. The first framework portrays the broader scope of the college education arena, moving from external factors to local factors. The second framework portrays the narrower range of the college classroom arena, including professor input and student input. The movement from a wide-ranging concern to a more limited concern illustrates how broadly worded public mandates distill into narrowly focused pedagogical practices.

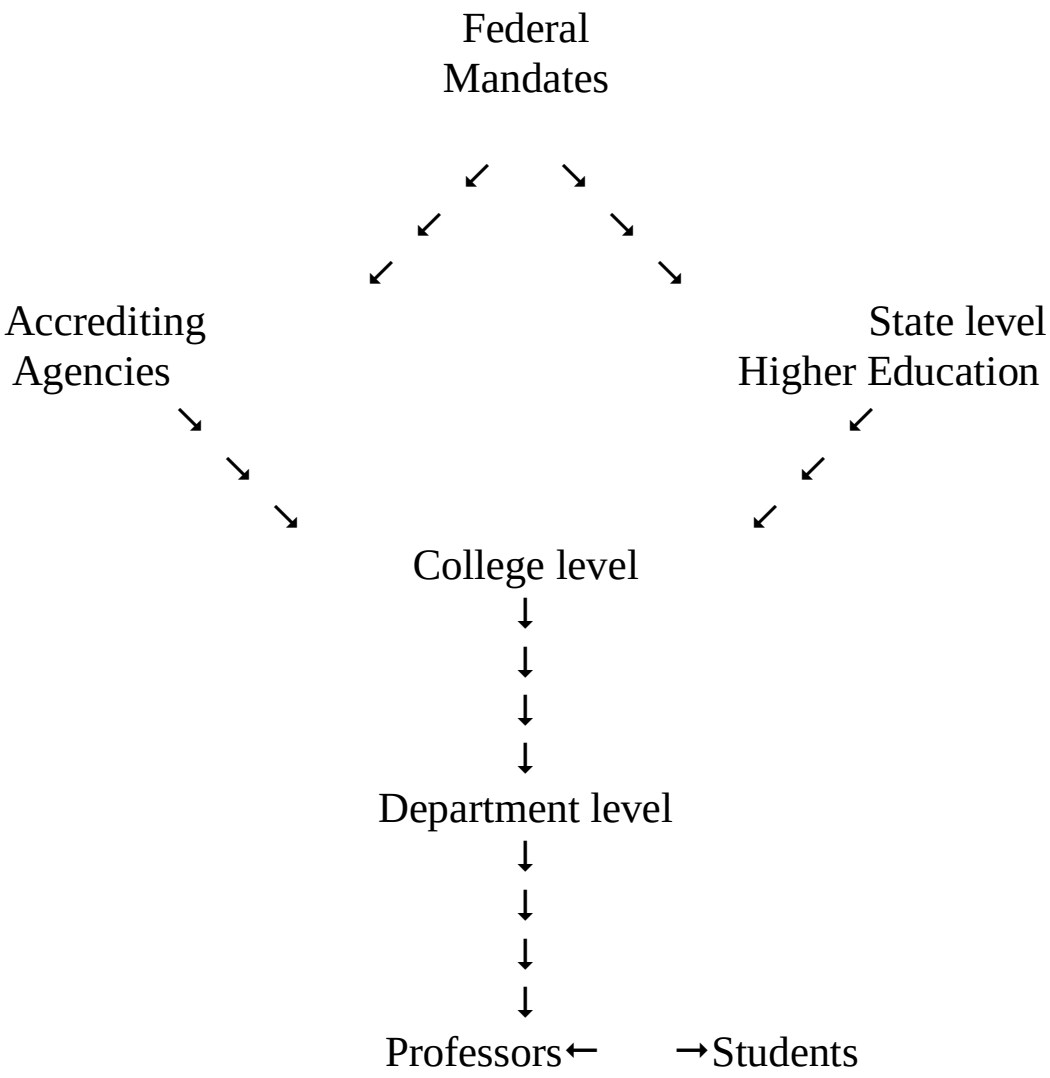
The first conceptual framework has illustrated how educational change has originated with federal mandates, then moves through the ranks of accreditation agencies and state boards of higher education. Colleges become the next echelon to act upon legislative mandates. Individual campus decisions are then passed on to departments and finally to professors in individual courses.

The second conceptual framework has diagrammed how professors have typically dealt with new directives. In order to plan a syllabus, they have considered the history, the theory, and the practices associated with the changes expected to impact the syllabus. These considerations have enabled professors to make decisions about classroom practices. In the college composition classroom, the assignments, essays, and measurements are all directed to the improvement of the writing skills and cognitive levels of students.

See both diagrams on the next two pages.

Educational Trends.

EXTERNAL FACTORS



LOCAL FACTORS

Local Factors.

PROFESSOR INPUT

Composition theory
+
Composition history
+
Composition practices:
Descriptive
Prescriptive
Cognitive

↙ ↙ ↙ ↙

Professor

↘ ↘ ↘

Classroom Practices:
assignments
essays
measurements

↘ ↘ ↘

Student:
Cognitive Skills
Writing Skills

STUDENT INPUT

Chapter Three: Methodology

This research study is quasi-experimental, nonequivalent control group design, an evaluative type of research. The subjects were assigned to predetermined groups (Writing I class sections) by the college registration system. It was the expectation that the students would show writing improvement as a result of completing this course. The goal of this research study was to investigate differences between the writing quality and cognitive level of research papers from the experimental group and the control group.

A Description of the Research Design

In the research study, ten different sections of freshmen college Writing I courses were included. Five courses, taught by the experimental professors, taught students using the Cognitive Level and Quality Writing Assessment (CLAQWA) rubric (see Appendices A and B). Five courses, taught by the control professors, did not teach students using the CLAQWA rubric.

The writing quality and cognitive level of the diagnostic/sample essays of students from the aforementioned courses were compared (see Appendices C2 and C4); the purpose of this comparison was to determine the equivalence of the two groups. The writing quality and cognitive level of the final/third essays of students from the aforementioned courses were compared. The purpose of this comparison was to investigate any statistically significant differences in the final essay scores of the two groups. Two independent raters scored the final essays according to the CLAQWA rubric. The researcher used the SPSS data analysis program to assess the scores. Tables, charts, and graphs illustrate the data analysis (see Table 3, see Figures 1 and 2). Commentary explains the meaning of the visual information.

The research site/demographics/context. The students in this research study all attend an urban state university in the Midwestern United States. The student body represents 46% male, 54% female, 19.6% minority, and 1% international students, and the campus has a population of over 14,000 undergraduate students. More than 90% of all students live in an area where the unemployment rate of 14% is above the state average unemployment rate of 10%. Over 90% of students commute to this state university campus from several neighboring counties.

This research study took place during two consecutive semesters in the university setting. Ten sections of Writing I courses participated. There were 55 students in the experimental group. The gender breakdown for the experimental group was 51% female and 49% male. There were 52 students in the control group. The gender breakdown for the control group was 48% female and 52% male. Both groups were similar demographically.

Ten freshmen writing sections of 25 students each, primarily 18-year-olds, were assigned to predetermined course sections via the campus registration system. This course, a Writing I course, is a general education requirement course for freshmen. The courses met for 50 minutes, 3 times per week or for 75 minutes, 2 times per week. Both met once weekly in a computer laboratory classroom setting and met the other times in a traditional classroom setting. The university requires all students to take this writing course, and the course focuses on quality writing and critical thinking skills (see Appendix D), both of which can be measured by the CLAQWA.

The general syllabus for the Writing I course indicates that students should be able to engage in critical thinking skills by analyzing arguments from readings and writing arguments of their own. Students should also be able to communicate effectively via the writing process.

Assignments in this course are geared to helping students gain these cognitive and writing skills, and many types of readings and writing activities are thought to help students meet this goal. Departmental guidelines suggest that students write at least three major writing assignments that are 1000-1200 words in length or 4-6 double-spaced pages, with the opportunity to revise their drafts so they produce a minimum of errors. The department expects all students to complete all assignments according to the schedule and policy of the instructor (see Appendix D).

The subjects/participants and how selected. The subjects of this research study were freshmen students enrolled at a Midwestern state university. The university requires all students to take a sequence of two freshman writing courses, and the first course of this sequence is under consideration for this research study. The first required writing course focuses on critical thinking skills and writing skills, both of which can me measured by the CLAQWA. Ten freshmen writing sections of 25 students each, primarily 18-year-olds, were assigned to predetermined groups for each course section. This course, a Writing I course, is a general education requirement course for freshmen. The course met for 50 minutes, 3 times per week or for 75 minutes, 2 times per week: once weekly in a computer laboratory classroom setting, other days in a traditional classroom setting.

The Instruments and Materials Used

The diagnostic/sample essay and the final/third essay, served as the data for this research study. Two raters evaluated them according to the CLAQWA instrument. Both the diagnostic and final essays from both the experimental group and the diagnostic and final essays from the control group were coded, so that raters did not know which papers belonged to which group. College students submitted assigned writing according to professor instructions. Although both the experimental group and the control group submitted other essays throughout the semester,

only the aforementioned essays were evaluated. In addition, only the essays of students who have turned in all previous writing assignments on time were used for this research. This type of selection represents the typical school situation (Ary, 2010, p. 316), and it accommodates the pretest/posttest quasi-experimental research design.

The research study examined differences between the writing skills and cognitive level of diagnostic essays and final essays written by college freshmenð between the experimental group, in which the professors implemented the CLAQWA rubric for instructional purposesð and the control group, in which the professors did not implement the same rubric.

The experimental professors implemented the CLAQWA in the following way. Before the assignments were given, the experimental professors explained the cognitive scale to the experimental group of 55 students, so that students were made aware of the aspects of thinking and learning that belong to each assignment (see Appendix A). Before each assignment was given, the experimental professors explained the parameters of the assignment, the writing skills, and other specific skills required in the assignment (Flateby & Metzger, 2001, p. 5). Professor instructions coincided with the guidelines of the CLAQWA rubric (see Appendix C). Professor instructions also coincided with departmental guidelines for the course (see Appendix D).

The control group did not implement the CLAQWA in the classroom. To the control group of 52 students, the control professors did not give the same directions. Nevertheless, professor instructions coincided with departmental guidelines for the course.

During the semester, the students in the experimental group submitted three essays to the experimental professors; these essays were evaluated according to the CLAQWA rubric. The diagnostic/sample essay was administered to students before instruction began; it was a timed, 50-minute essay. The third/final essay of the semester incorporated writing skills, structural

development, and analytical skills. The experimental professors evaluated student writing from all essays written during the semester according to criteria from the CLAQWA instrument.

During the semester, the students in the control group submitted three essays to the control professors; these essays were not evaluated according to the CLAQWA rubric. The diagnostic/sample essay was administered to students before instruction began; it was a timed, 50-minute essay. The third/final essay of the semester incorporated writing skills, structural development, and analytical skills. The control professors did not evaluate student writing from all essays written during the semester according to criteria from the CLAQWA instrument.

Analysis process. The analysis process began with three essays of a specified length, which are departmental requirements, so this comparison remains typical. The control professors neither taught nor evaluated student essays according to the CLAQWA rubric. The equivalence of both groups was measured by the diagnostic/sample essay. The performance of both groups was measured by the final/third essay.

All research papers were assessed by the CLAQWA, a 16-point rubric, based on a 2-scale system, an instrument designed to evaluate both writing quality and cognitive levels. Professors can use the scale separately or in combination for writing assignments. The cognitive levels of the 2-point scale are derived from the work of Bloom (1984). The CLAQWA groups the cognitive levels as follows: 1) knowledge, 2) comprehension, 3) application, and 4) analysis, synthesis, and evaluation. The writing quality assessment of the 2-point scale derives from commonly accepted goalsð like unity (Langan, p. 49; Lunsford, p. 20; Rosa & Eschholz, p. 90), support/development (Langan, p. 52; Lunsford, p. 20; Rosa & Eschholz, p. 164), coherence (Langan, p. 88; Lunsford, p. 21; Rosa & Eschholz, p. 190), and sentence skills (Langan, p. 366; Lunsford, p. 22-23; Rosa & Eschholz, p. 210-214),ð frequently named in college writing

textbooks. Both cognitive level and writing quality are evaluated on a 5-point continuum (Flateby & Metzger, n.d.).

The following tables are included after the reference section at the end of the dissertation: Table 1: Raw Data (SPSS Variables List); Table 2: Raw Data (SPSS Data File); Table 3: Summary of Results (Statistical Analysis from Steven McDonald, M.B.A.); Table 3.1: Inter rater reliability (Diagnostic Essay); Table 3.2: Inter Rater Reliability (Final Essay); Table 3.3: One Way ANOVA (Diagnostic Essay, experimental & control groups); Table 3.4: Word Skills Analysis (Final Essay, experimental & control groups); Table 3.5: Reasoning Skills Analysis (Final Essay, experimental & control groups); Table 3.6: One Way ANOVA for Final Essay (experimental & control groups); Table 3.7: One Way ANOVA (Diagnostic Essay, female & male); Table 3.8: One Way ANOVA (Final Essay, female & male).

The following figures are included after the tables section: Figure 1: Graphs of Scores (Control & Experimental Groups); Figure 2: Graphs of Scores (Female and Male Groups). The following appendices are included after the figures section: Appendix A: Cognitive Levels Chart; Appendix B: Cognitive Level and Quality Writing Assessment (CLAQWA) Rubric; Appendix C: Essay Assignments; Appendix C1: Diagnostic/Sample essay; Appendix C2: Diagnostic/sample essay rubric; Appendix C3: Final/third essay; Appendix C4: Final/third essay rubric; Appendix D: Departmental Rubric; Appendix E: Student Consent Form; Appendix F: Professor Permission Form; Appendix G: Permission from Dr. Terri Flateby; Appendix H: Permission from Youngstown State University; Appendix I: Permission from Liberty University.

By analyzing the diagnostic essays and final essays of 107 students, this researcher hypothesized that the students in the experimental group would show more significant outcomes

in overall writing skills and cognitive level than the control group. Before asking the research questions, inter rater reliability was ascertained from these two questions:

What is the difference between the scores from the two raters on the diagnostic essay?

What is the difference between the scores from the two raters on the final essay?

After determining the inter rater reliability, the following research questions were asked:

$RQ_{1:}$ What is the difference between the scores of the experimental group and the control group on the diagnostic essay?

$RQ_{2:}$ What effect does the use of the Cognitive Level and Quality Writing Assessment system (CLAQWA) instrument have on the writing skill of college students?

$RQ_{3:}$ What effect does the use of the Cognitive Level and Quality Writing Assessment system (CLAQWA) instrument have on the cognitive level of college students?

$RQ_{4:}$ What is the difference of scores between females and males on the diagnostic essay as graded by either rater?

$RQ_{5:}$ What is the difference of scores between females and males on the final essay as graded by either rater?

Selected points from the Cognitive Level and Quality of Writing Assessment (CLAQWA) instrument, a 16-point rubric, served as the measurement of writing quality and cognitive level. It was the expectation that students would show writing improvement in the culminating writing project of the semester.

In order to answer what initially seemed like a simple question, several hypotheses were investigated to arrive at a credible result. Inter rater reliability was determined before the comparison of final essay scores by looking at the significant difference between the scores from the two raters on the diagnostic essay and on the final essay. Gender differences were also

addressed. The following null hypotheses were stated:

H_1: There is no significant difference between the scores of the experimental group and the control group on the diagnostic essay.

H_2: There is no significant difference of word skill scores between the experimental group and the control group on the final essay.

H_3: There is no significant difference of cognitive level scores between the experimental group and the control group on the final essay.

H_4: There is no significant difference of scores between females and males on the diagnostic essay as graded by either rater.

H_5: There is no significant difference of scores between females and males on the final essay as graded by either rater.

The CLAQWA is a rubric developed at the University of South Florida (USF); it is an instrument intended ñto help instructors standardize their evaluation of writing and assess the cognitive level attained in student writingò (Flateby & Metzger, n.d., p. 2). Because colleges today understand that they face accountability and credibility issues in the public arena, colleges realize that such measurements can aid them in articulating their current and future goals.

The initial impetus for developing the rubric stemmed from the Universityôs stated mission: ñfacilitate the continuing improvement of é academic practices in the classroom and beyond,ò and ñassist degree programs [to] satisfy disciplinary, University-wide, and external criteria for excellenceò (Annual, 2006, p. 8). Many would agree that these goals are common to a wide spectrum of campuses.

According to USF, the campus recognized that many students were entering college inadequately prepared in the area of college writing skills. At the same time, the campus

understood that communication skills remain vital to the well-educated person. USF understood that good quality writing involves more than just correct sentence structure; good quality writing should also manifest growth in intellectual development (Annual, 2006, p. 8). Clearly, the campus needed to address this basic and essential concern.

The rubric has been in development since 1999 (Flateby & Metzger, 2001, p. 4), and it has helped both students and professors clarify writing objectives. The instrument was developed through the Office of Assessment to address issues related to student learning outcomes. The office formed committees to organize their assessment process. They conducted group and individual workshops for professors; they assessed learning outcomes and perceptions of learning experiences from students (Annual, 2006, p. 3). Writing skills and critical thinking skills retained the highest priority.

The Office of Assessment used the CLAQWA to assess student writing according to the following measurements: ñclass report from freshmen composition, summary report from all exit classes, comparison report using matched pairs of student essays é individual student reports, [and] the impact of systematic peer review on student writingò (Annual, 2006, p. 3). The evaluation of these measurements gave them the necessary tools to further refine the CLAQWA rubric. It has also undergone refinement and revision as a result of surveying faculty and students (Flateby, 2007).

The CLAQWA was developed with cross-disciplinary assessment of writing skills and cognitive levels in mind, and as such, it differs from other types of rubrics. The CLAQWA rubric is neither overly simplistic, like holistic scales nor overly complex, like rhetorical diagramming, yet it measures more than a dozen facets of writing (Annual, 2006). The design of the rubric allows the professor to tailor each assignment individually, so its apparent simplicity

augments its effectiveness. Such a rubric also allows campuses to communicate about student writing from a common framework. Because this rubric defines and standardizes the skills involved in the writing processes and outcomes, it allows a legitimate comparison between two sets of final papers. Departmental guidelines, like ñcritically analyze é develop their essaysò (see Appendix D) also help to make the comparison valid.

The researcher was situated in the complete observer role (Ary et al., 2010, p. 433). The experimental professors and control professors interacted with subjects (students) enough to establish rapport. The researcher merely collected the work activities of the subjects.

The data was evaluated by using the SPSS program to apply document analysis to written data. The main strength of this nonequivalent control group design is that the classes are already organized into class sections through the enrollment process. This frequently used quasi-experimental design works well in the college classroom because the student groups are already intact.

One concern to internal validity may be initial selection bias (Ary et al., 2010, p. 317), i.e. that important differences preexist between the two groups. This concern was controlled by administering a pretest to learn whether the groups are equivalent.

A second concern to internal validity may be experimenter effect (Ary et al., 2010, pp. 280-281), i.e. that students in the experimental group may respond more sensitively to the situation, knowing that they have consented to participate in a research design. This concern was controlled by letting professors (who are not the researcher) teach the course sections and by letting two other raters (not the professors of either group) evaluate the data independently.

Another concern to internal validity may be testing effect (Ary et al., 2010, pp. 274-275, 293), i.e. the knowing use of the pretest may cause the subjects to work more conscientiously

than they otherwise might have done. However, most would agree that, because students are accustomed to taking various forms of pretests, the use of the pretest will not significantly sensitize the subjects.

A fourth validity issue may be subject effects (Ary et al., 2010, pp. 281, 293). To control this issue, two independent raters evaluated the documents according to an established rubric called the Cognitive Level and Quality Writing Assessment (CLAQWA) system (see Appendix B).

A fifth validity issue may be attrition, sometimes a factor in freshmen general education requirement courses. It is true that only 55 out of 125 total students from the initial experimental group and 52 out of 125 total students from the initial control group remained to participate throughout the semester. However, the differential loss is so similar that attrition is not a serious threat. (Ary et al., 2010, p. 279).

A final validity issue may be the length of the study. Two semesters are not enough to speak with certainty about results, but two semesters are enough to suggest probable outcomes.

For the study, content analysis of written documents was conducted based on selected points from the Cognitive Level and Quality Writing Assessment (CLAQWA) instrument (Flateby & Metzger, 2001). The rubric is designed so that professors may select specific criteria according to assignment parameters. The rubric assessed change in the cognitive level and writing skills of college students from the experimental sections and the control sections, taught by different professors.

Each of the ten research groups was enrolled in a general education requirement course called Writing I. The enrolment limit for this freshmen course was 25 students per class, and

because Writing I is a general education requirement, enrolment was expected to be full in all course sections.

The English department gives various, general guidelines for the Writing I course, but each professor interprets the syllabus individually. During the semester, all professors assigned three writing assignments of similar length, according to departmental guidelines (see Appendix D). Departmental guidelines give some uniformity to the assignments in each section. Before the final assignment of the semester, both courses also likely included various types of other writing assignments, according to the discretion of the professor.

The experimental professors instructed the experimental group. The control professors taught the control group. The experimental group received feedback from the professors through the CLAQWA instrument about their skill levels in the three major essay assignments. The control group did not receive the same feedback from the other professors, but they did receive feedback not associated with CLAQWA.

Before instruction began, the experimental professors instructed the experimental classes about the cognitive scale (see Appendix A). The purpose of learning about cognition was to make students aware of the different purposes professors have when giving assignments. The experimental professors also showed students the CLAQWA rubric and explained its usefulness in assessing assignments consistently.

Before the diagnostic/sample essay was assigned to the experimental group, the experimental professors explained the parameters of the assignment (see Appendices C1 and C2). From the CLAQWA rubric, two independent raters evaluated student writing according to the following CLAQWA traits:

- trait 2 - main idea

- trait 7 - paragraph construction
- trait 8 ï closing
- trait 16 - grammar and mechanics
- trait 1 - assignment requirements.

Throughout the semester, the experimental professor explained other assignments according to the CLAQWA rubric and evaluated student essays according to selected CLAQWA traits.

Before the final/third essay was assigned to the experimental group, the experimental professors explained the requirements of the assignment (see Appendices C3 and C4). From the CLAQWA rubric, the experimental professors evaluated student writing according to several CLAQWA traits:

- trait 5 ï opening
- trait 6 ï coherence devices
- trait 8 ï closing
- trait 9 ï reasoning
- trait 10 ï quality of details
- trait 12 ï word choice
- trait 13 ï comprehensibility.

Students turned in a draft of the assignment after receiving instruction. The experimental professors marked the essays according the CLAQWA rubric; students had the opportunity to edit their writing and resubmit for a higher grade, a common practice in college writing courses.

Two experienced college English faculty members served as independent raters of both the diagnostic/sample essays and the final/third papers, according to the CLAQWA rubric. All

papers were coded so that the raters did know which papers belonged to either group. Following here is a title list of the two essay assignments that were used:

- Diagnostic/Sample Essay and CLAQWA classroom rubric (see Appendices C1 and C2)
- Final/Third Essay and CLAQWA classroom rubric (see Appendices C3 and C4)

The data analyses made. The document analysis was performed using the SPSS statistics program to assess the scores from the CLAQWA instrument on the cognitive level and writing quality of diagnostic/sample essays and final/third essays. The charts from the program were incorporated into the experimental report. This study investigated any statistically significant differences in the final essay scores of 107 first year college students: 55 in the experimental group and 52 in the control group. The analysis of scores consisted of several statistical measurements.

To simplify the scoring process, the researcher set up the SPSS program to analyze the following types of information:

1. The SPSS variables list, which names the variables (see Table 1)
2. The SPSS data file, which includes the summary statistics from the four sets of scores (see Table 2)

The researcher used the SPSS program to analyze the following data and answer the corresponding questions:

1. the reliability of the two ratersô scores to each other (see Tables 3.1 and 3.2)

   What is the difference between the scores from the two raters on the diagnostic essay?

   What is the difference between the scores from the two raters on the final essay?
2. the homogeneity/equivalence of the two groups to each other (see see Table 3.3)

$RQ_{1:}$ What is the difference between the scores of the experimental group and the control group on the diagnostic essay?

3. the difference in word skill scores of the final essay between the two groups (see Table 3.4)

   $RQ_{2:}$ What effect does the use of the Cognitive Level and Quality Writing Assessment system (CLAQWA) instrument have on the writing skill of college students?

4. the difference in cognitive level scores of the final essay between the two groups (see Table 3.5)

   $RQ_{3:}$ What effect does the use of the Cognitive Level and Quality Writing Assessment system (CLAQWA) instrument have on the cognitive level of college students?

5. the difference in total scores of the final essay between the two groups (see Table 3.6)

   This is addressed separately in RQ_2 and $RQ_{3.}$

6. the difference in male/female scores of both groups (see Tables 3.7 and 3.8)

   $RQ_{4:}$ What is the difference of scores between females and males on the diagnostic essay as graded by either rater?

   $RQ_{5:}$ What is the difference of scores between females and males on the final essay as graded by either rater?

The following measurements were used:

1. To determine the reliability between the two raters of the student essays, the following measurement was used:

- The Pearson Correlation gave the reliability between the two raters by using the mean scores from the both groups or two variables (Ary et al., 2010, p. 256).
    - This measurement allows the researcher to make accurate comparisons between groups.
    - This measurement indicates whether or not the results are reliable and valid.

2. To determine the homogeneity/equivalence of scores between the two groups (pretest/diagnostic essay), the following measurement was used:
    - The One Way ANOVA gives the homogeneity between the two groups (Ary et al., 2010, p. 178).
        - This measurement tests the null hypothesis that there is no difference between the two groups.
3. To establish the difference in word skill scores of the final essay between the two groups, the following measurement was used:
    - The One Way ANOVA gives the difference between the word skill scores of the final essays of the 2 groups (Ary et al., 2010, p. 178).
        - This measurement indicates whether the students in the experimental group showed greater gains in word skill scores at the end of the semester than the control group.
        - This measurement tests the null hypothesis that no difference existed between the two groups for word skill score.
4. To determine the difference in cognitive level/reasoning skill scores of the final essay between the two groups, the following measurement was used:

- The One Way ANOVA gives the difference between the cognitive level scores of the final essays of the 2 groups (Ary et al., 2010, p. 178).
    - This measurement indicates whether the students in the experimental group showed greater gains in cognitive level scores at the end of the semester than the control group.
    - This measurement tests the null hypothesis that no difference existed between the two groups for cognitive level scores.

5. To establish the difference in total/combined scores of the final essay between the two groups, the following measurement was used:
    - The One Way ANOVA gives the difference between the scores of the final essays of the 2 groups (Ary et al., 2010, p. 178).
        - This measurement indicates whether the students in the experimental group showed greater gains at the end of the semester than the control group.
        - This measurement tests the null hypothesis that no difference existed between the two groups.
6. To indicate differences between male and female scores, the following measurement was used:
    - The One Way ANOVA gives the difference between male and female scores on the final essays of the two groups (Ary et al., 2010, p. 178).
        - This measurement may suggest whether gender affects outcomes.
7. To portray a visual representation of the final score differences, graphs were used at the end of the data section.

CLAQWA RUBRIC

8. To summarize mean data from the above measurements, a chart has been included at the end of the data section.

This proposal has been approved by the Human Subjects Research Committee (HSRC) at Youngstown State University. This proposal has been approved by the Institutional Review Board (IRB) at Liberty University.

A Summary Statement of the Methodology

In order to learn whether students from the experimental group scored significantly higher than students from the control group, the research study employed a nonequivalent control group design. Scores from the experimental group were compared to scores from the control group. The Cognitive Level and Quality Writing (CLAQWA) was the assessment instrument. Two independent raters scored the essays according to selected criteria from the 16-point rubric. This researcher used SPSS, a data analysis program to assess the scores. This research study investigated any statistically significant differences in the final/third essay scores.

Chapter Four: Results of the Study

The results of this study about college composition are organized according to the following pattern. First, chapter four introduces the research problem and explains the research methodology. Second, this chapter summarizes the research results. Third, the appendices give the permission data and the human subject research permissions. Finally, the raw data obtained from the research is listed, and the summary statistics and explanations from the statistician are included.

Presentation of the Results

These results describe the differences of cognitive level and writing quality of college students enrolled in ten sections of Writing I courses. By analyzing the diagnostic essays and final essays of 107 students, the researcher hypothesized that the students in the experimental group would show more significant outcomes in overall writing skills and cognitive level than the control group. The purpose of this comparison was to see whether the use of the specific rubric resulted in higher scores (of either word skill or reasoning skill) on the final essays of the experimental group when compared to the control group. The data was analyzed using the SPSS program.

Two independent raters scored the same group of essays, using the same rubric for grading guidelines. Before determining differences between the two groups of essays, the scores from both raters were compared for statistically equivalent correlation. Then, the scores of both groups were compared for homogeneity. Because the campus description of the student body reports minorities as one category, ethnicity was not addressed here, but for exploratory analysis,

female/male differences were compared for changes in scores on the final essays. Finally, tables and graphs illustrate the results of this research study.

Research Problems/Problem Statements

The project investigated two primary research statements: 1) College students who are instructed from the CLAQWA rubric will score higher in writing skill on final essays than students instructed by an alternative method; and 2) College students who are instructed from the CLAQWA rubric will score higher in cognitive level on final essays than students instructed by an alternative method. In order to lend credibility to the primary research statements, inter rater reliability was first ascertained from these two questions:

What is the difference between the scores from the two raters on the diagnostic essay?

What is the difference between the scores from the two raters on the final essay?

After determining the inter rater reliability, the following research questions were asked:

$RQ_{1:}$ What is the difference between the scores of the experimental group and the control group on the diagnostic essay?

$RQ_{2:}$ What effect does the use of the Cognitive Level and Quality Writing Assessment system (CLAQWA) instrument have on the writing skill of college students?

$RQ_{3:}$ What effect does the use of the Cognitive Level and Quality Writing Assessment system (CLAQWA) instrument have on the cognitive level of college students?

$RQ_{4:}$ What is the difference of scores between females and males on the diagnostic essay as graded by either rater?

$RQ_{5:}$ What is the difference of scores between females and males on the final essay as graded by either rater?

This research assumed that college students, whose writing was evaluated by the CLAQWA instrument, would score higher on their final papers than college students who were not instructed by the same method. Several null hypotheses were tested here:

H_1: There is no significant difference between the scores of the experimental group and the control group on the diagnostic essay.

H_2: There is no significant difference of word skill scores between the experimental group and the control group on the final essay.

H_3: There is no significant difference of cognitive level scores between the experimental group and the control group on the final essay.

H_4: There is no significant difference of scores between females and males on the diagnostic essay as graded by either rater.

H_5: There is no significant difference of scores between females and males on the final essay as graded by either rater.

The null hypotheses were tested by the use of the CLAQWA rubric, an instrument that measures growth in cognitive development as well as writing skills, so the use of this rubric allowed the problem statement to combine both measurements into one question. (see Appendix B).

Professional Significance of the Problem

Initiated by the No Child Left Behind legislation at the beginning of the twenty-first century, colleges are now faced with academic accountability, and this means that post-secondary institutions are increasingly concerned with demonstrating that their students have shown measurable progress in the courses they take. Because freshmen writing courses are mandated at nearly every undergraduate institution, the sequential, freshmen writing courses have become a logical focus of such measurement.

Research about freshman college writing is important because colleges face tremendous pressure to validate their course offerings. So much has been written about every facet of the college writing curriculum that the selection of related information has become an ongoing task for anyone who wishes to investigate even one aspect of student writing. Studies that shed light on whether or not a specific method is better than another method can offer a much needed perspective on how to think through this new requirement about accountability.

Overview of the Methodology

This quasi-experimental, evaluative research study was a nonequivalent control group design. The subjects were assigned to predetermined groups (Writing I class sections). All of the students wrote frequently, and they were expected to show improvement in the final essay as a result of completing the course.

In this research study, college freshmen from ten different sections of Writing I courses participated. Five courses, taught by the experimental professors, taught and evaluated students using the CLAQWA rubric. Five courses, taught by the control professors, neither taught nor evaluated students using the CLAQWA rubric.

The writing quality and cognitive level quality of the final papers of the two groups of students were compared by using the Cognitive Level and Quality Writing Assessment (CLAQWA) rubric. Two independent raters scored the final papers according to the CLAQWA rubric. This researcher used the SPSS data analysis program to assess the scores. This research study investigated any statistically significant differences in the final paper scores. The purpose of this comparison was to determine whether teaching writing according to a specific rubric yielded greater improvement (in either word skill or reasoning skill) in student essays than not teaching writing according to the specific rubric.

The subjects/participants and how selected: The subjects of this research study were freshmen students enrolled at an urban Midwestern state university. The university required all students to take a sequence of two freshman writing courses, and the first course of this sequence was under consideration for this research study. The first required writing course focused on critical thinking skills and writing skills, both of which can be measured by the CLAQWA. Among ten freshmen writing sections of 25 students in each section (students who are primarily 18-year-olds), students were assigned to predetermined course sections via the college enrollment process. This course, a Writing I course, is a general education requirement course for freshmen. The course met for three clock hours per week, three times weekly or twice weekly: once weekly in a computer laboratory classroom setting, the other class time in a traditional classroom setting.

The instruments and materials used: The diagnostic/sample essay and the final/third essay, served as the data for this research study. Two raters evaluated the student essays according to the CLAQWA instrument. Both the diagnostic and the final essays from both the experimental group and the control group were coded, so that raters did not know which papers belonged to which group. College students submitted assigned writing according to professor instructions. Although both the experimental group and the control group submitted other essays throughout the semester, only the aforementioned essays were evaluated. In addition, only the essays of students who turned in all previous writing assignments on time were used for this research. This typical case sampling is necessary for an accurate comparison.

This research study examined differences between the writing skills and cognitive level of diagnostic and final essays written by college freshmenð between the experimental group, in

which the professors implemented the CLAQWA rubric for instructional purposesð and the control group, in which the professors did not instruct from the same rubric.

The purpose of comparing the diagnostic essays was to determine whether the two groups showed any significant differences between them, in other words, whether the experimental group and the control group could be considered homogeneous, or similar in skill level, at the beginning of the semester. The purpose of comparing the final essays was to determine whether the two groups showed any significant differences in either word skill outcomes or reasoning skill outcomes, in other words, whether the experimental group showed greater skill gains than the control group at the end of the semester.

The experimental professors implemented the CLAQWA in the following way. Before the assignments were given, the experimental professors explained the use of the cognitive scale (from the Cognitive Levels chart) to the experimental group of 55 students (see Appendix A), so that students were made aware of the aspects of thinking and learning that take place when a professor gives each assignment. Before each assignment was given, the experimental professors explained the parameters of the assignment, the writing skills, and other specific skills required in the assignment. Instructions from the professor coincided with the guidelines of the CLAQWA rubric (see Appendix B). The experimental students received feedback from the CLAQWA rubric.

The control professors did not did not implement the CLAQWA in the classroom. Before the assignments were given, the control professors did not explain the use of the cognitive scale. To the control group of 52 students, the control professors did not give the same instructions. The control students did not receive feedback from the CLAQWA rubric.

CLAQWA RUBRIC

By analyzing 107 studentsô diagnostic essays and final essays, the researcher hypothesized that the students in the experimental group would show more significant outcomes in overall writing skills and cognitive level than the students in the control group did. Several null hypotheses have already been stated.

During the semester, the students in the experimental group submitted three essays to the experimental professors; these essays were evaluated according to the CLAQWA rubric. The diagnostic/sample essay was administered to students before instruction began; it was a timed, 50-minute essay. The third/final essay of the semester was the culminating effort of the semester; it incorporated writing skills, structural development, and analytical skills. The experimental professors evaluated student writing from all essays written during the semester according to criteria from the CLAQWA instrument.

During the semester, the students in the control group submitted three essays to the control professors; these essays were not evaluated according to the CLAQWA rubric. The diagnostic/sample essay was administered to students before instruction began; it was a timed, 50-minute essay. The third/final essay of the semester was the culminating effort of the semester; it incorporated writing skills, structural development, and analytical skills. The control professors did not evaluate student writing from any essays written during the semester according to criteria from the CLAQWA instrument.

Three essays of a specific length are departmental requirements, so this comparison remains typical. The control professors neither taught nor evaluated student essays according to the CLAQWA rubric. The equivalence of both groups was measured by the diagnostic/sample essay. The performance of both groups was measured by the final/third essay. Performance was separated into word skill scores and reasoning skill scores.

CLAQWA RUBRIC

College students submitted writing assignments according to the instructions of each professor. Although both the experimental group and the control group submitted other essays throughout the semester, only the diagnostic essay and the final essay were evaluated. In addition, only the essays of students who turned in all previous writing assignments on time were used for this research. This is a typical case sampling; it is assumed that students who turn in all of the previous work of the semester are typical, not atypical.

All final/third essays papers were assessed by the CLAQWA, a 16-point rubric that can evaluate both writing skills and cognitive skills. Professors can use the scale values (writing, cognitive) separately or in combination, to meet specific needs of specific writing assignments. The cognitive values of knowledge, comprehension, application, and analysis/synthesis/evaluation were derived from the well-known work of Benjamin S. Bloom (1984). The writing quality values were derived from commonly accepted goals, like unity (Langan, p. 49; Lunsford, p. 20; Rosa & Eschholz, p. 90), support/development (Langan, p. 52; Lunsford, p. 20; Rosa & Eschholz, p. 164), coherence (Langan, p. 88; Lunsford, p. 21; Rosa & Eschholz, p. 190), and sentence skills (Langan, p. 366; Lunsford, p. 22-23; Rosa & Eschholz, p. 210-214), goals frequently named in college writing textbooks. Both writing quality skills and cognitive level skills were evaluated on a 5-point continuum.

The quasi-experimental, nonequivalent control group design was used for the research study. Two types of scores from the experimental group were compared to two types of scores from the control group. The Cognitive Level and Quality Writing (CLAQWA) was the assessment instrument. Two independent raters scored the essays according to selected criteria from the 16-point rubric. The researcher used SPSS, a data analysis program, to assess the

scores for statistically significant differences in the final essay scores, tallied separately into word skill scores and reasoning skill scores.

Summary in General Terms of the Results Obtained

To measure overall change in writing quality, an experimental group and a control group submitted essays. The experimental group was instructed from a specific rubric, and the control group was not instructed from a specific rubric.

Each of the 107 students submitted two essays: a diagnostic essay before instruction began, at the beginning of the semester; and a final essay after completing instruction, toward the end of the semester. The two raters scored each essay according to the CLAQWA rubric (see Appendices B, C2, and C4).

Sample Sizes of Groups

Experimental group	55 participants
Control group	52 participants
	107 total participants

The statistical tables on the following pages were written by the statistician, Steven McDonald. The analyses on the following pages were written by the writer of this dissertation.

To determine inter rater reliability: There is no significant difference between the scores from the two raters on the diagnostic essay.

Table 3.1
Inter rater reliability for Diagnostic Essay

Descriptive Statistics

	Mean	Std. Deviation	N
diag1	16.9252	4.90610	107
diag2	14.4953	3.39637	107

Correlations

		diag1	diag2
diag1	Pearson Correlation	1	.636**
	Sig. (2-tailed)		.000
	N	107	107
diag2	Pearson Correlation	.636**	1
	Sig. (2-tailed)	.000	
	N	107	107

**. Correlation is significant at the 0.01 level (2-tailed).

To test the null hypothesis of homogeneity of variance between the two raters, the Pearson correlation was run on the diagnostic essay scores. On the diagnostic essay, the mean score from the first rater was 16.9252, with a standard deviation of 4.90610. On the diagnostic essay, the mean score from the second rater was 14.4953, with a standard deviation of 3.39637. The Pearson correlation was .636 (sig. = 0.00) at Ŭ= 0.01 in a two-tailed test.

To determine inter rater reliability: There is no significant difference between the scores from the two raters on the final essay.

Table 3.2
Inter Rater Reliability for Final Essay

Descriptive Statistics

	Mean	Std. Deviation	N
final1	25.5421	6.33541	107
final2	24.7850	3.09604	107

Correlations

		final1	final2
final1	Pearson Correlation	1	.252**
	Sig. (2-tailed)		.009
	N	107	107
final2	Pearson Correlation	.252**	1
	Sig. (2-tailed)	.009	
	N	107	107

**. Correlation is significant at the 0.01 level (2-tailed).

To test the null hypothesis of homogeneity of variance between the two raters, the Pearson correlation was run on the final essay scores.

On the final essay, the mean score from the first rater was 25.5421, with a standard deviation of 6.33541. On the final essay, the mean score from the second rater was 24.7850, with a standard deviation of 3.09604. The Pearson correlation was .252 (sig. = 0.00) at Ŭ= 0.01 in a two-tailed test.

The test confirmed that both raters scored similarly. Results suggested that both raters showed a statistically significant correlation in grading both the diagnostic essay and the final essay. Similar scoring indicates that the validity of the results is reliable (see Table 3.1).

H_1: There is no significant difference between the scores of the experimental group and the control group on the diagnostic essay.

Table 3.3
One Way ANOVA for Diagnostic Essay (experimental and control groups)

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
diag1	experimental	55	17.0182	4.60471	.62090	15.7734	18.2630	6.00	25.00
	control	52	16.8269	5.24954	.72798	15.3654	18.2884	5.00	25.00
	Total	107	16.9252	4.90610	.47429	15.9849	17.8656	5.00	25.00
diag2	experimental	55	14.2364	3.10891	.41921	13.3959	15.0768	8.00	20.00
	control	52	14.7692	3.68684	.51127	13.7428	15.7957	8.00	25.00
	Total	107	14.4953	3.39637	.32834	13.8444	15.1463	8.00	25.00

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
diag1	.900	1	105	.345
diag2	1.077	1	105	.302

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
diag1	Between Groups	.978	1	.978	.040	.841
	Within Groups	2550.424	105	24.290		
	Total	2551.402	106			
diag2	Between Groups	7.590	1	7.590	.656	.420
	Within Groups	1215.158	105	11.573		
	Total	1222.748	106			

To test the null hypothesis that there is no significant difference between the experimental group and the control group in performance on the diagnostic essay, the One Way ANOVA was run on the diagnostic essay scores.

The test indicated that the groups were homogeneous. The mean score from the diagnostic essays of the experimental group was 17.0182, with a standard deviation of 4.60471. The mean score from the diagnostic essays of the control group was 16.8269, with a standard deviation of 5.24954. The Levene statistic tests suggested that no statistically significant differences were shown between the two groups (see Table 3.3).

A lack of difference here indicates that both groups begin at a similar level at the beginning of the semester, implying that measureable change at the end of the semester will give an accurate measurement of change.

H_2: There is no significant difference of word skill scores between the experimental group and the control group on the final essay.

Table 3.4
Word Skills Analysis for Final Essay (experimental and control groups)

Descriptives

		N	Mean	Std. Deviation	Std. Error
wordsk1	experimental	55	10.4545	2.03505	.27441
	control	52	10.6923	1.52802	.21190
	Total	107	10.5701	1.80205	.17421
wordsk2	experimental	55	14.1455	2.02227	.27268
	control	52	14.2308	1.42272	.19730
	Total	107	14.1869	1.74902	.16908

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
wordsk1	3.855	1	105	.052
wordsk2	4.244	1	105	.042

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
wordsk1	Between Groups	1.511	1	1.511	.463	.498
	Within Groups	342.713	105	3.264		
	Total	344.224	106			
wordsk2	Between Groups	.195	1	.195	.063	.802
	Within Groups	324.067	105	3.086		
	Total	324.262	106			

To test the null hypothesis that there is no significant difference of word skill scores between the experimental group and the control group, the One Way ANOVA was run on the final essays.

The test suggested that the scores of both groups were similar. On the final essays, the mean of the word skill scores of the experimental group was 10.4545 from rater one and 14.1455 from rater two, with an average mean of10.5701 and a standard deviation of 1.80205. The Levene statistic from both raters showed little significant difference between mean word skill scores for the experimental group and the control group.

A lack of difference here suggests that both the experimental group and the control group made similar gains in word skill scores (see Table 3.4), indicating that the experimental treatment differed little from the control treatment.

H_3: There is no significant difference of cognitive level scores between the experimental group and the control group on the final essay.

Table 3.5
Reasoning Skills Analysis for Final Essay (experimental and control groups)

Descriptives

		N	Mean	Std. Deviation	Std. Error
reason1	experimental	55	11.1455	2.84422	.38351
	control	52	10.7115	2.61471	.36260
	Total	107	10.9346	2.73092	.26401
reason2	experimental	55	14.9273	3.74094	.50443
	control	52	14.2692	3.54310	.49134
	Total	107	14.6075	3.64396	.35227

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
reason1	.826	1	105	.366
reason2	.171	1	105	.680

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
reason1	Between Groups	5.033	1	5.033	.673	.414
	Within Groups	785.509	105	7.481		
	Total	790.542	106			
reason2	Between Groups	11.574	1	11.574	.871	.353
	Within Groups	1395.940	105	13.295		
	Total	1407.514	106			

To test the null hypothesis that there is no significant difference of cognitive level scores between the experimental group and the control group, the One Way ANOVA was run on the final essays.

On the final essay, the mean of the reasoning skill scores of the experimental group was 11.1455 from rater one and 14.9273 from rater two, with an average mean of 10.9346 and a standard deviation of 2.73092.

The Levene statistics from both raters showed little significant difference between mean reasoning skill scores for the experimental group and the control group. The One Way ANOVA failed to reject the null hypothesis (see Table 3.5).

A lack of difference here indicates that both groups made similar gains in cognitive level scores, suggesting that the experimental treatment varied little from the control treatment.

H_4: There is no significant difference of scores between females and males on the diagnostic essay as graded by either rater.

Table 3.7
One Way ANOVA for Diagnostic Essay (female and male)

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
diag1	female	52	16.5192	5.00388	.69391	15.1261	17.9123	5.00	25.00
	male	55	17.3091	4.82614	.65076	16.0044	18.6138	5.00	25.00
	Total	107	16.9252	4.90610	.47429	15.9849	17.8656	5.00	25.00
diag2	female	52	14.0769	3.25303	.45111	13.1713	14.9826	8.00	25.00
	male	55	14.8909	3.51016	.47331	13.9420	15.8398	8.00	22.00
	Total	107	14.4953	3.39637	.32834	13.8444	15.1463	8.00	25.00

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
diag1	.000	1	105	.993
diag2	.708	1	105	.402

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
diag1	Between Groups	16.676	1	16.676	.691	.408
	Within Groups	2534.726	105	24.140		
	Total	2551.402	106			
diag2	Between Groups	17.710	1	17.710	1.543	.217
	Within Groups	1205.038	105	11.477		
	Total	1222.748	106			

.3091 for males, with an average score of 16.9252 and a standard deviation of 4.90610. The

To test the null hypothesis of no significant difference between female scores and male scores on the diagnostic essay, the One Way ANOVA was run.

The mean score of the diagnostic essay scores from the rater one was 16.5192 for females and 17mean score of the diagnostic essay scores from the rater two was 14.0769 for females and 14.8909 for males, with an average score of 14.4953 and a standard deviation of 3.39637. The Levene statistic from both raters showed no significant differences between genders on this score.

These results indicated that no statistically significant differences were shown between females and males on the diagnostic essay from either rater (see Table 3.7). The gender differentiated scores suggest that both genders began the treatment similarly.

H_5: There is no significant difference of scores between females and males on the final essay as graded by either rater

Table 3.8
One Way ANOVA for Final Essay (female and male)

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
final1	female	52	26.8462	5.78160	.80176	25.2365	28.4558	14.00	35.00
	male	55	24.3091	6.63569	.89476	22.5152	26.1030	7.00	35.00
	Total	107	25.5421	6.33541	.61247	24.3278	26.7563	7.00	35.00
final2	female	52	25.2500	3.21074	.44525	24.3561	26.1439	16.00	33.00
	male	55	24.3455	2.94529	.39714	23.5492	25.1417	15.00	30.00
	Total	107	24.7850	3.09604	.29931	24.1916	25.3784	15.00	33.00

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
final1	1.187	1	105	.278
final2	.510	1	105	.477

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
final1	Between Groups	172.046	1	172.046	4.425	.038
	Within Groups	4082.515	105	38.881		
	Total	4254.561	106			
final2	Between Groups	21.870	1	21.870	2.310	.132
	Within Groups	994.186	105	9.468		
	Total	1016.056	106			

To test the null hypothesis of no significant difference between female scores and male scores on the final essay, the One Way ANOVA was run.

The mean score of the final essay scores from the rater one was 26.8462 for females and 24.3091 for males, with an average score of 25.5421 and a standard deviation of 6.33541. The mean score of the final essay scores from rater two was 25.2500 for females and 24.3455 for males, with an average score of 24.7850 and a standard deviation of 3.09604.

The Levene statistic indicated that a significant difference existed between the female and male scores from rater one, but not from rater two. This difference may prove to be interesting, and it may warrant further analysis as composition studies may look at gender differences expressed in writing outcomes (See Table 3.8).

The results suggested a statistically significant difference between females and males from the score of rater one, but not from the scores of rater two. Females scored slightly higher than males from rater one scores. This difference may open up an avenue of pursuit for researchers who wish to investigate gender differences in the outcomes of a research project because the inter rater reliability for both graders is consistent.

The graphs of scores for both groups are shown in Figure 1 and Figure 2. First, the graphs of scores for the experimental group and the control group for both essays are shown. Second, the graphs of scores for the female group and male group for both essays are shown.

It is hoped that this research study will make a contribution to the implementation of changes within individual college writing classrooms. While many anecdotal studies report on diverse methods of success, none have yet addressed the specific requirements of accountability arising from recent legislation like the Action Plan for Higher Education from the Commission on the Future of Higher Education.

Chapter Five: Discussions and Conclusions

Chapter 5 summarizes and discusses the research findings and correlates the findings to seminal research. The results of this study about college composition offer a perspective on teaching to the college community, particularly to English departments, in the area of freshmen college writing. Finally, chapter 5 offers an explanation concerning the limitations of the study and addresses the need for future research.

Summary and Discussion of Findings

The statistical analysis of this data does not support the initial assumption of the research studyð that the experimental group of students who were taught according to a specific rubric would show greater improvement in writing skills and reasoning skills than the control group who was not taught according to the same rubric. Although the experimental group showed minimal improvement over the control group in the combined scores, the improvement shown was not statistically significant. Both groups showed similar enough improvement in writing skills and cognitive skills that it could not be definitely stated that the improvement could be attributed to the use of the specific rubric. Instead, the data supports a widely understood idea that numerous teaching methods and styles can accomplish similar results, an important concept for departments to bear in mind as campuses approach the task of measuring change in student writing.

This conclusion may seem more acceptable to some because it supports the many and varied approaches of professors to teach in the way that seems best to them. The statistical analysis of this data may also mean that the control professors were already teaching a mixture of

methods that may have included elements of prescriptive, descriptive, and cognitive methods (Duncan, 2007), i.e. a combination that includes the more balanced or environmental ways of teaching writing. The statistical analysis of this data may also mean that other, carefully presented rubrics may accomplish similar ends.

In contrast to what was expected, the statistical analyses of the data confirmed a commonly held assumption about freshmen college students, that ñmost students should show improvementò (Hillocks, 1995, p. 207). In other words, professors can feel confident that their manner of explaining a given assignment to students will enable them to succeed at writing.

This type of measurement gives one answer to a simple question that has been generated as a result of recent legislative mandates. Because most English departments give broad guidelines for course syllabi, teaching style preferences are not at issue here. Yet, quantitative measurements of writing change may be beneficial to campuses as they strive to meet educational mandates.

This research should remind educators that rating essays according to a rubric does not perform the same function as grading an essay. The purpose of grading an essay is specialized and geared to individual students. Writing change is expected from students as a result of grading. The purpose of rating according to a rubric is generic and geared to established standards. Writing change has already occurred at the time of rating. Rubrics that are intended for public purposes can indicate improvement, but they cannot be compared to the copious grading that professors give to student essays, nor can they be compared to the very exacting standards of standardized testing measurements that students and professionals may be familiar with.

This research indicates that professors can continue to teach to realistic goals commonly understood to be the aims of most college level writing courses and can continue to expect to see student improvement without worrying about teaching according to a specific method. In addition, colleges need not move in the direction of standardized teaching methods, for educational freedom remains of paramount importance to instructors. At the same time, colleges are advised to learn how to conduct educational research appropriately, with the understanding that such research is limited in what it can measure.

Clearly, rubrics will continue to hold central importance to those who evaluate student writing for personal, institutional, or accountability purposes. To return to the advice of Hillocks, departments and professors must think carefully about what they are doing and how they are doing it (1995, p. 31). They must also continue with determination to clearly classify and state their intended goals, most likely through the use of a rubric (Hillocks, 1986, Chapter 8).

Using an appropriate rubric will allow professors the freedom in teaching that is so central to any academic discussion. The Cognitive Level and Quality Writing Assessment (CLAQWA) instrument was developed specifically with college level learning in mind. Although the CLAQWA was initially developed for faculty use, its revisions quickly included institutional objectives and student functions. One may expect that just as federal education policy leads at the regional and state level, so will regional and state bodies lead departments and professors to similar ends. College professors, who have typically had little training in the area of education, may be expected to clarify their course level goals more articulately. Perhaps departments may employ this rubric as a teaching tool for instructors. The quantification of results and consistent record keeping may allow for more long-range comparisons, beneficial to departments when conducting self-assessment.

The use of the CLAQWA instrument gives all professors in all disciplines a common ground for talking about student outcomes, and its importance to academic discussions about assessment cannot be overlooked. Although this research did not indicate that first-year college writing showed a statistically significant difference (i.e. improvement) through the use of the CLAQWA instrument, the results do not in any way negate the efficacy of the rubric. The results merely suggest that the rubric may already be closely allied to common writing goals and may also suggest that English professors typically teach this way as a matter of common practice. The use of this specific rubric will still enable campuses to further implement institutional goals that are now coming to the attention of English departments from legislative and accreditation sources. The use of a standardized rubric in conjunction with departmental goals may enable departments to fulfill legislative directives more completely and efficiently.

The shift toward academic accountability occurring in American education is both extensive and colossal. It is the beginning of a new orientation and emphasis in education at all levels that has yet to be fully implemented. An obvious feature of academic accountability is the measurement of outcomes. Both the steadfast emphasis on critical thinking skills and the logical assumption that student writing reflects cognitive levels have contributed to the way educators perceive accountability. While many instruments measure many aspects of all types of outcomes, it is of the utmost importance to determine whether or not standardized instruments that reflect common understandings of writing skills and cognitive levels are indeed necessary to the implementation of what seems to be the trend in American undergraduate education. This study may begin to answer this question; nevertheless, many more studies, using other rubrics, may also contribute to answering the same question.

This research confirms that guidance from professorsð neither a complete domination of instruction nor a total absence of instructionð helps students learn the skills of the course. This coincides with the findings of several researchers (Chickering & Gamson, 1987; Cooper & Odell, 1977; Davis, 1989, 1990; Duncan, 2007; Hillocks, 1986, 1995) who have inquired, in a systematic way, about writing practices, habits, and methods since the 1980s up to the current time.

While this research study confirms that students do improve their skills during the duration of the semester-long course with or without the use of a specific rubric, it in no way undermines the efficacy of the CLAQWA rubric, an instrument that measures both cognitive level and quality of writing in college student essays. Nevertheless, this research study does affirm that the use of a specific rubric may not be the universal answer to measurement concerns. As a result, this research supports the diverse initiatives of professors to teach according to their preference.

One assumption behind such legislation is that there is a better way or a right way to accomplish the task of measuring improvement in student writing. Many colleges have begun to work on identifying and clarifying these objectives. Clearly, this legislation does not necessarily reflect the generalized message of research in compositionð that appropriate classroom management can foster the cognitive growth that reveals itself through writing. Legislation can make colleges view the answer to the question of measurement as existing outside of their own expertise concerning their own student bodies and departmental programming. Contrary to legislative assumptions, this research refutes a prevailing idea that there is one specific way to teach writing or one specific way to measure change in writing. Instead, it supports previous

research about composition, research that comes from pedagogical, historical, psychological, and quantitative perspectives, research that began in the 1980s and continues to the present day.

At the same time, the analyses raised points that may be of interest to future researchers, such as the relationship of gender to writing outcomes, the nature of the subject area, or what professors may reasonably expect to accomplish when teaching composition. Because gender differences are beginning to be investigated at other levels of education, female/male differences may also warrant further investigation at the college level.

Correlation to Seminal Research

Hillocks (1986) offered the most comprehensive and outstanding coverage of composition studies. His initial effort actually covered more than two decades of research material, already outdated by several years by the time he published his findings. Nevertheless, he made several points very clear. First, that out of the more than 500 articles on composition studies, only 73, or about 14%, were suitable for the type of quantitative research he was trying to summarize (p. 187). This low percentage reveals several flaws seen in many composition studies: lack of researcher control for teacher bias, lack of researcher control for rater bias, lack of the use of pre-writing and post-writing samples, and lack of rating compositions for quality (p. 134). Composition studies may also attest to the diverse nature of composition and the diverse ways of evaluating it.

The work of Hillocks indicated that quantification may become more important to the academic area of college writing in order to talk about the subject (Hillocks, 1986, Chapter 8), and this study contributes a quantitative study that may be of use to future collators of composition studies. It appears that while such efforts have begun, a more intense focus may be valuable for future discussions.

Second, the work of Hillocks (1995) emphasized the importance of self-awareness of (teaching) actions, and its corollary, ñthe option to continue or discontinue themò (p. 189), an emphasis that continues to this day and seems now to be mandated by educational reform legislation. In other words, teachers must continually evaluate and revise what they have done in the classroom if meaningful, long-term change is to take place. This serious and time-consuming task, labeled as reflective practice (p. 28), is essential to accurate measurement and assessment of student outcomes

Duncan (2007) provided a valuable and logical starting point to renew collegiate interest in composition studies. By identifying the historical trends of composition teaching and by specifying three major ways of teaching college composition (descriptive, prescriptive, cognitive), Duncan has allowed the multifarious writing methods/programs/strategies to be categorized in a clearly understandable way. This succinct way of looking at composition studies is necessary to meaningful communication about the subject.

It would be impractical to underestimate the advantages of many classroom studies, whose significance hold tremendous value for classroom practitioners, and no one would like to see such studies disappear from publication in educational circles. A need exists for accessible information, but recent legislation seems to insist that more research, of a different type, be promoted. This study belongs among studies that shift toward such quantitative studies in writing. Rubrics can be designed to measure both cognitive skills and quality of writing, so studies that employ a tailored rubric may contribute to quantitative studies.

Theoretical implications of the study. One aspect of the specific requirements of accountability arising from recent legislation like the Action Plan for Higher Education from the Commission on the Future of Higher Education (*Action Plan* 2006) concerns the way that

student work will be perceived as a result of its enactment and enforcement. The legislation seems to be asking for quantitative measurements by focusing on specifics that may direct the efforts of both instructors and students.

This researcher understands that different ways of thinking from one academic discipline to another constitute a greater barrier to the understanding and implementation of legislation than any other factor. To put it another way, whereas English professors generally think of a piece of student writing as an intellectual product, legislation requiring the measurement of writing outcomes changes the piece of student writing from an intellectual product into an aspect of human behavior that occurs in the classroomð because the student does the act of writing the essay. Understandably, this shifts the focus of writing away from the perspective of English toward the perspectives of Psychology (human behavior) and Education (occurs within a classroom). In addition to the shift from the Arts to the Social Sciences, the most important factor in measuring human behavior is human will, which cannot be measured quantitatively.

Because cognition is a relatively new area of study within the educational arena, perhaps additional research could focus on aspects of cognition that appear difficult to define and measure. It may be that professors, department heads, and researchers would require training in multiple disciplines in order to conduct their investigations into these cognitive aspects of learning.

Limitations and unanticipated findings. One limitation of this research study concerns the choice of final papers that were considered for research purposes. Only the final papers of students who submitted all previous writing assignments on time were considered in this research study. Another limitation of the proposed study may be attrition, sometimes a factor in college student attendance, and certainly a factor in this study. Any course or study cannot

measure improvement in all students, if all students do not complete the course. A third limitation of the research study is the length of the study. Two semesters are not enough to speak with certainty about results, but two semesters are enough to suggest probable outcomes.

Research did not corroborate the primary assumption that the experimental group of students would show greater improvement than the control group of students in their final essays. Results indicated no significant differences in performance between the two groups, either in the word skill score, the reasoning skill score, or the combined mean score. Because the statistical analysis showed that both groups were considered to be homogeneous and that the scores of both raters were considered to be comparable to one another, this outcome must be explained some other way.

The absence of significant difference may be due, in part, to the similarity of the use of the CLAQWA rubric to the typical teaching procedures of freshmen writing courses. The experimental professors found the rubric so adaptable to their personal style of teaching, that experimental and control groups may have been more alike than different. This suggests that a different way to compare outcomes may be in order.

Second, only about half of the original students who formally consented to participate in the research project actually completed the semester-long course. It may be that only the stronger academic students were measured. Perhaps these student outcomes indicated a narrower range of improvement.

The study did not inquire into studentsô composing habits but focused instead on professorsô outcomes. This suggests that studying other factors, which may also affect writing outcomes, may enhance the understanding of student outcomes, and as a result, how to bring about maximum improvement.

Implications for Practice

While many anecdotal studies report on diverse methods of success, the time is ripe to address the specific requirements of accountability arising from recent legislation like the Action Plan for Higher Education from the Commission on the Future of Higher Education (*Action Plan*, 2006). The results of this study raised several points of interest to the college writing community. The results may:

- Contribute to quantitative composition studies by acknowledging prior research and its link to current goals
- Underscore the nature of composition measurements and how composition is perceived by various disciplines
- Support the importance of research about writing as a means of clarifying terminology and disseminating standards
- Emphasize the need for more research into cognitive aspects of writing as a way to fulfill current legislative aims
- Recognize how rubrics can aid instruction and establish common goals
- Affirm individual teaching styles and understand how personal practices can relate to current goals

This research study will broaden the understanding of professors and departments as they implement changes that arise from legislation concerning measurement of outcomes in college writing classrooms.

Future Research and Final Thoughts

Because the cognitive aspects of learning affect the teaching of writing, more research into how cognition and writing skill are related should be investigated. It is commonly

understood that students have different learning styles as much as it is understood that professors possess different teaching styles, and that subject areas or courses require different ways of teaching. The meaning of cognition and its purposes within the undergraduate classroom need to be disseminated to college professors.

One interesting result from this study concerned the area of how gender differences affect learning. Because one rater showed a difference between male and female scores (but the other rater did not), gender differences at the college level may be an area for further investigation. Just as gender differences have begun to be studied in the early elementary grades, so may gender differences continue to be investigated up through the college level of learning.

Finally, individual departments should continue or begin to engage in collaborative sessions where professors evaluate student essays, course syllabi, and grading rubrics. The ongoing discussions about how to teach and what to teach generally prove invaluable to individual instructors, but the discussions also serve a broader function of keeping departmental effort consistent for the purpose of consistent measurement.

References

Action plan for higher education: Improving accessibility, affordability and accountability. (2006). Retrieved from United States Department of Education, No Child Left Behind page: http://www.ed.gov/nclb

Annual report of academic assessment activities: Academic year 2005-2006. (2006). Retrieved from University of South Florida, Assessment Center: http://usfweb2.usf.edu/assessment/index.shtml

Ary, D., Jacobs, L.C., & Sorensen, C. (2010). *Introduction to research in education* (8th ed.). Belmont, CA: Thomson Wadsworth.

Bloom, B.S., ed. (1984). *Taxonomy of educational objectives: The classification of educational goals, Handbook 1, Cognitive domain*. New York: Longman.

Boards and Commissions: Higher education update. (2006). Retrieved from United States Department of Education, No Child Left Behind page: http://www.ed.gov/nclb

Cazort, D. (1982). Advice from a recent has-been to a TA starting cold. *Freshman English News, 10*(3), 1-4.

Chickering, A.W., & Gamson, Z.F. (1987). Seven principles for good practice in undergraduate education. *American Association for Higher Education Bulletin*, *39*(7), 3-7.

CLAQWA online: A new tool for improving student writing and thinking. (2007). Retrieved from University of South Florida, Assessment Center: http://usfweb2.usf.edu/assessment/index.shtml

CLAQWA online: Cognitive level and quality of writing assessment. (2006). Retrieved from University of South Florida, Assessment Center: http://usfweb2.usf.edu/assessment/index.shtml

Cooper, C.R., & Odell, L. (Eds.). (1977). Holistic evaluation of writing. In C.R. Cooper and L. Odell (Eds.), *Evaluating Writing* (pp. 3-31). Urbana, IL: National Council of Teachers of English.

Corbett, E.P.J., & Connors, R.J. (1999). *Classical rhetoric for the modern student*. (4th ed.). New York: Oxford University Press.

Cross, K.P. (1999). What do we know about studentsô learning, and how do we know it? *Innovative Higher Education, 23*(4), 255-270.

Dave, A.N., & Russell, D.R. (2010). Drafting and revision using word processing by undergraduate student writers: Changing conceptions and practices. *Research in the Teaching of English, 44*(4), 406-434.

Davidson-Shivers, G.V., Nowlin, B., & Lanouette, M. (2002). Do multimedia lesson structure and learning styles influence undergraduate writing performance?. *College Student Journal*, *36*(1), 20-32.

Davis, W.K. (1989). *The effects of Christensen's 'generative rhetoric of the sentence' on the overall quality and number of complex sentences in college freshmen's writing.* ERIC Document Reproduction Service No. ED491490.

Davis, W.K. (1990). *The effects of process-centered and form-centered instruction on the coherence of college freshman writing*. ERIC Document Reproduction Service No. ED316878.

DuBoulay, D. (1999). Argument in reading: What does it involve and how can students become better critical readers? *Teaching in Higher Education, 4*(2), 147-163.

Duncan, M. (2007). Whatever happened to the paragraph? *College English, 69*(5), 470-495.

Dwyer, E.J. (1992). Using a journalism model for teaching essay writing. *Journal of Reading, 36*(3), 226-232.

Eades, C. (2002). A working model of pedagogical triangulation: A holistic approach to peer-revision workshops. *Teaching English in the Two-Year College, 30*(1), 60-67.

Elder, L. (2000). Critical thinking: Nine strategies for everyday life, part II. *Journal of Developmental Education, 24*(2), 38-40.

Elder, L., & Paul, R. (2001). Critical thinking: Thinking to some purpose. *Journal of Developmental Education, 25*(1), 40-42.

Elder, L. & Paul, R. (2002). Critical thinking: Distinguishing between inferences and assumptions. *Journal of Developmental Education, 25*(3), 34-36.

Executive summary of changes to the principles of accreditation. (2007). Retrieved from Southern Association of Colleges and Schools, Commission on Colleges: www.sacscoc.org

Flateby, T.L. (2007). *Closing the loop: Assessing and improving writing and thinking throughout the curriculum*. Retrieved from University of South Florida, Assessment Center: http://usfweb2.usf.edu/assessment/index.shtml

Flateby, T.L., & Eubanks, D. (2008). *CLAQWA online: Next steps in writing assessment*. Retrieved from University of South Florida, Assessment Center: http://usfweb2.usf.edu/assessment/index.shtml

Flateby, T.L., & Metzger, E. (2001). Instructional implications of the cognitive level and quality of writing assessment (CLAQWA). *Assessment Update, 13*(1), 4-5,11.

Flateby, T.L., & Metzger, E.A. (n.d.). *CLAQWA: Cognitive level and quality writing assessment system*. Retrieved from University of South Florida, Assessment Center: http://usfweb2.usf.edu/assessment/index.shtml

Ford, D.G. (2002). Teaching anecdotally. *College Teaching, 50*(3), 114-116.

Four pillars of NCLB. (2004). Retrieved from United States Department of Education, No Child Left Behind page: http://www.ed.gov/nclb

Friend, R. (2001). Effects of strategy instruction on summary writing of college students. *Contemporary Educational Psychology, 26*, 3-24.

Hafer, G.R. (2001). Ideas in practice: Supplemental instruction in freshman composition. *Journal of Developmental Education, 24*(3), 30-36.

Hillocks, G. (1986). *Research in written composition: New directions for teaching*. Urbana, IL: National Conference of Research in English.

Hillocks, G, (1995). *Teaching writing as reflective practice*. New York: Teachers College Press.

Horning, A.S. (1997). Reflection and revision: Intimacy in college writing. ERIC Document Reproduction Service No. ED403564.

Huot, B. (2007). Consistently inconsistent: Business and the Spellings Commission report on higher education. *College English, 69*(5), 512-525.

Jacobs, S.E. (1977). A study in coherence: Applications for the teaching of writing, teacher training, and research. ERIC Document Reproduction Service No. ED145432.

Langan, J. (2006). *English skills with readings* (6th ed.). Boston: McGraw Hill.

Lavelle, E., & Zuercher, N. (2001). The writing approaches of university students. *Higher Education, 42*(3), 373-391.

Lunsford, A.A. (2010). *Easy writer: A pocket reference* (4th ed.). Boston: Bedford/St. Martinôs.

Marzano, R.J. (1983). Toward a working model for the analysis of cohesion. ERIC Document Reproduction Service No. ED241930.

Mohammed, A., & Jaber, H. (2008). The effects of deductive and inductive approaches of teaching on Jordanian university studentsô use of the active and passive voice in English. *College Student Journal, 42*(2), 545-553.

Oldenburg, S. (2006). Grammar in the student-centered composition class. *Radical Teacher, 75*, 43.

Olson, L. (2005). Purpose of testing needs to shift, experts say. *Education Week, 25*(8), 7-9.

Peer review of student writing examined. (2006). Retrieved from University of South Florida, Assessment Center: http://usfweb2.usf.edu/assessment/index.shtml

Perry, W.G. (1999). *Forms of intellectual and ethical development in the college years: A scheme*. San Francisco: Jossey-Bass.

Primary trait analysis used by the CLAQWA system. (2007). Retrieved from University of South Florida, Assessment Center: http://usfweb2.usf.edu/assessment/index.shtml

Ramey, S., VandeVusse, L., & Gosline, M. (2007). Using a written communication rubric to improve studentsô writing. *International Journal of Learning, 13*(10), 67-74.

Roen, D.H. (1984). Coherence in writing. *English Journal, 73*(5), 35-36.

Roen, D.H., Pantoja, V., Yena, L., Miller, S.K., & Waggoner, E. (Eds.). (2002). Strategies for teaching first-year composition. ERIC Document Reproduction Service No. ED459476.

Rosa, A. & Eschholz, P. (2007). *Models for writers: Short essays for composition* (9th ed.). Boston: Bedford/St. Martinôs.

Sadler, B., & Andrade, H. (2004). The writing rubric. *Educational Leadership, 62*(2), 48-53.

Sanders, L.R. (2001). Improving assessment in university classrooms. *College Teaching, 49*(2), 62-65.

Sanders, S.E., & Littlefield, J.H. (1975). Perhaps test essays can reflect significant improvement in freshman composition: Report on a successful attempt. *Research in the Teaching of English, 9*(2), 145-153.

Shelley, J.R. (1998). Incorporating computer literacy into the composition classroom. ERIC Document Reproduction Service No. ED422584.

Shuman, R.B. (1991). What are the priorities? *Educational Leadership, 49*(2), 82.

Soles, D. (2005). An analysis of the style of exemplary first-year writing. *Teaching English in the Two Year College*, *33*(1), 38-49.

U.S. Department of Education awards $2.45 million for improved measures of achievement for post-secondary students. (2007). Retrieved from United States Department of Education, No Child Left Behind page: http://www.ed.gov/nclb

USF Assessment philosophy statement. (2007). Retrieved from University of South Florida, Assessment Center: http://usfweb2.usf.edu/assessment/index.shtml

Westscott, W., & Ramey, B. (1993). Aiming for excellence in freshman English. ERIC Document Reproduction Service No. ED361697.

Williams, J.M. (1997). *Style: Ten lessons in clarity and grace*. (5th ed.). New York: Addison Wesley Longman Inc.

Williams, J.M. (1990). *Style: Toward clarity and grace*. Chicago: University of Chicago Press.

Writing and thinking assessment: Revealing problems and offering solutions. (2007). Retrieved from University of South Florida, Assessment Center: http://usfweb2.usf.edu/assessment/index.shtml

Youngstown State University. (2007). *Undergraduate bulletin 2007-2008*. Youngstown, OH: Youngstown State University.

Appendix A
Cognitive Levels Chart

	Rating
Level 4: Analysis, Synthesis, Evaluation • Makes a judgment of a work or plan based upon a given or constructed set of specific criteria, not opinion. • Organizes or reorganizes ideas or combines elements to make a whole. • Distinguishes between fact and fiction. • Compares and contrasts or deduces. • Identifies relationships of parts to the whole.	**High 5** **Medium 3** **Low 1**
Level 3: Application • Uses what is understood in a new situation. • Uses what is learned in the assignment or in class.	**High 5** **Medium 3** **Low 1**
Level 2: Comprehension • Translates or rephrases known words, interprets or explains in a way that demonstrates understanding of the material.	**High 5** **Medium 3** **Low 1**
Level 1: Knowledge • Accurately recalls or describes, or identifies information which was presented in class or reading. • Involves memorization.	**High 5** **Medium 3** **Low 1**

Appendix B
Cognitive Level and Quality Writing Assessment (CLAQWA) Rubric

ASSIGNMENT PARAMETERS

Level	Trait 1: Assignment Requirements
5	The writer addresses and develops each aspect of the assignment. (Quality is judged through other categories.)
4	The writer addresses each aspect of the assignment.
3	The writer addresses the appropriate topic and partially fulfills assignment requirements.
2	The writer addresses the appropriate topic, but omits most or all of the assignment requirements.
1	The writer is off topic or vaguely addresses the topic.

Level	Trait 2: Main Idea
5	The writer clearly has and maintains a main idea throughout.
4	The main idea is clear, although a rare extraneous element is introduced.
3	The paper has a main idea, but additional unrelated ideas distract the reader.
2	The main idea is not maintained or it is unclear.
1	The paper lacks a main idea or appears to reflect the writer's "free association."

Level	Trait 3: Audience
5	The writer exhibits a keen awareness of the audience's needs and expectations.
4	The writer exhibits an awareness of the audience's needs and expectations.
3	The writer exhibits reader awareness and addresses the appropriate audience throughout the text, although in some sections the audience is ambiguous.
2	The writer shows a lack of reader awareness by addressing one or more inappropriate audiences.
1	The writer shifts between multiple and/or inappropriate audiences because of a lack of reader awareness.

Level	Trait 4: Purpose
5	The elements of the paper clearly contribute to the writer's purpose, which is obvious, specific, maintained, and appropriate for the assignment.
4	The writer's purpose is present, appropriate for the assignment, and maintained throughout.
3	The writer's purpose is present and appropriate for the assignment, but elements may not clearly contribute to the purpose.
2	The writer presents multiple purposes or the purpose is inappropriate for the assignment.
1	The writer's purpose is not evident.

ORGANIZATION AND DEVELOPMENT: STRUCTURAL INTEGRITY

Level	Trait 5: Opening
5	The writer uses the opening to introduce the main idea, capture the reader's attention, and prepare the reader for the body of the paper.
4	The writer uses the opening to introduce the main idea and prepares the reader for the body of the paper.
3	The writer uses the opening to identify the main idea, but does not prepare the reader for the body of the paper.
2	The main idea is not clear from the opening.
1	The opening is absent or is unrelated to the main idea.

Level	Trait 6: Coherence Devices
5	Transitional words, phrases, sentences and paragraphs (coherence devices) smoothly connect the paper's elements, ideas and/or details, allowing the reader to follow the writer's points effortlessly.
4	Coherence devices are rarely missing and do not impact the reader's understanding.
3	Coherence devices appear throughout the paper, but additional and appropriate connectors would enhance the flow.
2	Coherence devices are attempted but are ineffective.
1	Coherence devices are absent or inappropriate.

Level	Trait 7: Paragraph Construction
5	Each paragraph is unified around a topic that relates to the main idea. All paragraphs support the main idea and are ordered logically.
4	Paragraphs support the main idea and are ordered logically, but an occasional paragraph may not be unified around a single topic.
3	Paragraphs exist but some may be misplaced, include more than one topic, or be unrelated to the main idea.
2	Paragraph breaks are attempted but are illogical and misplaced. Topics may also be unrelated to the main idea.
1	There are no paragraph breaks. Topics may be unrelated to the main idea and presented illogically.

Level	Trait 8: Closing
5	Closing synthesizes the elements, supports the main idea, and finalizes the paper.
4	Closing summarizes the elements, supports the main idea, and finalizes the paper.
3	Closing summarizes the elements, supports the main idea, may introduce unrelated or new details, but does not finalize the paper.
2	Closing presents a few elements which are consistent with the main idea, may introduce unrelated or new ideas, but does not finalize the paper.
1	Closing is absent or introduces unrelated ideas.

ORGANIZATION AND DEVELOPMENT: REASONING & CONSISTENCY

Level	Trait 9: Reasoning
5	The essay exhibits a logical progression of sophisticated ideas that support the focus of the paper.
4	The essay exhibits a logical progression of ideas that support the focus of the paper.
3	The progression of ideas is interrupted by rare errors in logic, such as absolutes or contradictions.
2	The attempt at a progression of ideas is unsuccessful due to errors in logic, such as absolutes or contradictions.
1	The ideas are illogical and appear to reflect the writer's "stream of consciousness."

Level	Trait 10: Quality of Details
5	Details help to develop each element of the text and provide supporting statements, evidence or examples necessary to explain or persuade effectively.
4	Details support the elements of the text with sufficient clarity, depth and accuracy.
3	Details are related to the elements of the text, but do not support those elements with sufficient clarity, depth and accuracy.
2	Details are loosely related to the elements of the text, but are lacking clarity, depth and accuracy.
1	Details do not develop the elements of the text.

Level	Trait 11: Quantity of Details
5	All points are supported by a sufficient number of details.
4	Most points are supported by a sufficient number of details.
3	Additional details are needed to develop **some** points.
2	Additional details are needed to develop **most** points.
1	Virtually no details are present.

GRAMMAR AND MECHANICS: OBSERVATION OF STANDARD EDITED ENGLISH

Level	Trait 16: Grammar and Mechanics
5	Sentences are grammatically and mechanically correct.
4	Rare grammatical and mechanical errors exist, but do not affect readability.
3	A limited variety of grammatical errors exist.
2	A variety of grammatical errors appear throughout the paper possibly affecting readability.
1	Most sentences exhibit multiple grammatical and mechanical errors, obstructing meaning.

LANGUAGE: CONTEXTUAL AND AUDIENCE APPROPRIATENESS

Level	Trait 12: Word Choice
5	Vocabulary reflects a thorough grasp of the language appropriate to the audience. Word choice is precise, creating a vivid image. Metaphors and other such devices may be used to create nuanced meaning.
4	Vocabulary reflects a strong grasp of the language appropriate to the audience. Word choice is accurate.
3	Vocabulary reflects an inconsistent grasp of the language and may be inaccurate or inappropriate to the audience.
2	Vocabulary is typically inaccurate and inappropriate to the audience. Word choice may include vague, non-descriptive, and/or trite expressions.
1	Word choice is limited to vague, non-descriptive, and/or trite expressions and may include homonyms, errors, word choice inappropriate to the audience, and "thesaurus writing."

Level	Trait 13: Comprehensibility
5	All sentences are clear and understandable.
4	The sentences are clear and understandable with rare ambiguities.
3	Most sentences are understandable but may include ambiguities.
2	Many sentences lack clarity and may misuse academic language.
1	Most sentences lack clarity and may misuse academic language.

Level	Trait 14: Sentence Construction
5	Clear and concise sentences vary, with the degree of complexity reflecting the audience and purpose.
4	Sentences vary, with the degree of complexity reflecting the audience and purpose.
3	Sentence variety is limited but attempts complex structure.
2	Complex structure is attempted without success and/or sentence structure is simplistic, but not throughout the text.
1	Sentences are simple and repetitive.

Level	Trait 15: Point of View
5	Point of view is consistent and appropriate for the purpose and audience.
4	Point of view is appropriate for the purpose and audience, and a rare shift returns to the original point of view.
3	Point of view shifts occasionally, or may be consistent but inappropriate, for the purpose and/or audience.
2	Point of view is attempted, but shifts frequently.
1	Point of view is not established, confusing the reader.

Appendix C
Essay Assignments

Appendix C1
Diagnostic/Sample Essay

The diagnostic/sample essay is graded on the following scale:

- Main idea
- Paragraph Construction
- Closing
- Grammar and Mechanics
- Assignment parameters
    - This refers to an overall evaluation of the writing
    - The topics may vary, but the topic is not being evaluated as much as the writing

Appendix C2
Diagnostic/Sample Essay Rubric

RUBRIC:DIAGNOSTIC/SAMPLE ESSAY

Level	Trait 2: Main Idea
5	The writer clearly has and maintains a main idea throughout.
4	The main idea is clear, although a rare extraneous element is introduced.
3	The paper has a main idea, but additional unrelated ideas distract the reader.
2	The main idea is not maintained or it is unclear.
1	The paper lacks a main idea or appears to reflect the writer's "free association."

Level	Trait 7: Paragraph Construction
5	Each paragraph is unified around a topic that relates to the main idea. All paragraphs support the main idea and are ordered logically.
4	Paragraphs support the main idea and are ordered logically, but an occasional paragraph may not be unified around a single topic.
3	Paragraphs exist but some may be misplaced, include more than one topic, or be unrelated to the main idea.
2	Paragraph breaks are attempted but are illogical and misplaced. Topics may also be unrelated to the main idea.
1	There are no paragraph breaks. Topics may be unrelated to the main idea and presented illogically.

Level	Trait 8: Closing
5	Closing synthesizes the elements, supports the main idea, and finalizes the paper.
4	Closing summarizes the elements, supports the main idea, and finalizes the paper.
3	Closing summarizes the elements, supports the main idea, may introduce unrelated or new details, but does not finalize the paper.
2	Closing presents a few elements which are consistent with the main idea, may introduce unrelated or new ideas, but does not finalize the paper.
1	Closing is absent or introduces unrelated ideas.

Level	Trait 16: Grammar and Mechanics
5	Sentences are grammatically and mechanically correct.
4	Rare grammatical and mechanical errors exist, but do not affect readability.
3	A limited variety of grammatical errors exist.
2	A variety of grammatical errors appear throughout the paper possibly affecting readability.
1	Most sentences exhibit multiple grammatical and mechanical errors, obstructing meaning.

Level	Trait 1: Assignment Requirements
5	The writer addresses and develops each aspect of the assignment. (Quality is judged through other categories.)
4	The writer addresses each aspect of the assignment.
3	The writer addresses the appropriate topic and partially fulfills assignment requirements.
2	The writer addresses the appropriate topic, but omits most or all of the assignment requirements.
1	The writer is off topic or vaguely addresses the topic.

Appendix C3
Final/Third Essay

The final/third essay is graded on the following scale:

- Opening
- Coherence Devices
- Closing
- Reasoning
- Quality of Details
- Word Choice
- Comprehensibility

Appendix C4
Final/Third Essay Rubric

Level	Trait 5: Opening
5	The writer uses the opening to introduce the main idea, capture the reader's attention, and prepare the reader for the body of the paper.
4	The writer uses the opening to introduce the main idea and prepares the reader for the body of the paper.
3	The writer uses the opening to identify the main idea, but does not prepare the reader for the body of the paper.
2	The main idea is not clear from the opening.
1	The opening is absent or is unrelated to the main idea.

Level	Trait 6: Coherence Devices
5	Transitional words, phrases, sentences and paragraphs (coherence devices) smoothly connect the paper's elements, ideas and/or details, allowing the reader to follow the writer's points effortlessly.
4	Coherence devices are rarely missing and do not impact the reader's understanding.
3	Coherence devices appear throughout the paper, but additional and appropriate connectors would enhance the flow.
2	Coherence devices are attempted but are ineffective.
1	Coherence devices are absent or inappropriate.

Level	Trait 8: Closing
5	Closing synthesizes the elements, supports the main idea, and finalizes the paper.
4	Closing summarizes the elements, supports the main idea, and finalizes the paper.
3	Closing summarizes the elements, supports the main idea, may introduce unrelated or new details, but does not finalize the paper.
2	Closing presents a few elements which are consistent with the main idea, may introduce unrelated or new ideas, but does not finalize the paper.
1	Closing is absent or introduces unrelated ideas.

Level	Trait 9: Reasoning
5	The essay exhibits a logical progression of sophisticated ideas that support the focus of the paper.
4	The essay exhibits a logical progression of ideas that support the focus of the paper.
3	The progression of ideas is interrupted by rare errors in logic, such as absolutes or contradictions.
2	The attempt at a progression of ideas is unsuccessful due to errors in logic, such as absolutes or contradictions.
1	The ideas are illogical and appear to reflect the writer's "stream of consciousness."

Level	Trait 10: Quality of Details
5	Details help to develop each element of the text and provide supporting statements, evidence or examples necessary to explain or persuade effectively.
4	Details support the elements of the text with sufficient clarity, depth and accuracy.
3	Details are related to the elements of the text, but do not support those elements with sufficient clarity, depth and accuracy.
2	Details are loosely related to the elements of the text, but are lacking clarity, depth and accuracy.
1	Details do not develop the elements of the text.

Level	Trait 12: Word Choice
5	Vocabulary reflects a thorough grasp of the language appropriate to the audience. Word choice is precise, creating a vivid image. Metaphors and other such devices may be used to create nuanced meaning.
4	Vocabulary reflects a strong grasp of the language appropriate to the audience. Word choice is accurate.
3	Vocabulary reflects an inconsistent grasp of the language and may be inaccurate or inappropriate to the audience.
2	Vocabulary is typically inaccurate and inappropriate to the audience. Word choice may include vague, non-descriptive, and/or trite expressions.
1	Word choice is limited to vague, non-descriptive, and/or trite expressions and may include homonyms, errors, word choice inappropriate to the audience, and "thesaurus writing."

Level	Trait 13: Comprehensibility
5	All sentences are clear and understandable.
4	The sentences are clear and understandable with rare ambiguities.
3	Most sentences are understandable but may include ambiguities.
2	Many sentences lack clarity and may misuse academic language.
1	Most sentences lack clarity and may misuse academic language.

Appendix D
Departmental Rubric

ENGLISH 1550: Writing 1

English Department General Syllabus

I. Course Description (from the 2006-2007 *Undergraduate Student Bulletin*)

"Strategies for writing as a means of critical inquiry, with focus on writing processes and on the roles of writer, audience, and purpose as they affect writing. Students divide their time between regular classrooms and computer classrooms, where they have the opportunity to acquire and develop basic word-processing and electronic communication skills. Open to students on the basis of Composition and Reading Test results or successful completion of ENGL 1539 or ENGL 1540. Grading is ABC/NC. 3 s.h." ***Writing 1 is a General Education requirement.***

A. Course focus

Writing 1 aims to help students cultivate college-level reading, writing, and thinking abilities, in a way that will serve them well as educated citizens and as students at YSU.

B. Placement

Placement into Writing 1 is based on successful completion of English 1539 or 1540, or Composition Placement Test results along with ACT English sub-scores, SAT Verbal sub-scores, and/or COMPASS™ Reading Test Score.

C. Fee

YSU requires student to pay a technology and materials fee in Writing 1. This fee is used to maintain and replace equipment, software, and supplies.

D. Students in Writing 1 will:

- Read, discuss, and critically analyze primarily nonfiction prose; other readings may be given as supplements.
- Write primarily nonfiction, expository essays; other writing assignments may be given as supplements.
- Develop their essays through the use of multiple drafts, peer reviews, and instructor comments.
- Use computer labs and other online resources as tools for cultivating their writing.

E. YSU Syllabus Requirements

Per YSU/YSU-OEA *Agreement*, Article 25.3, all syllabi must include a grading policy and an attendance policy.

II. Texts

Unless otherwise noted, the most recent editions of textbooks will be used. Required texts in English 1550 include *The Little, Brown Compact Handbook* (6th ed.) and one of the following textbooks (according to the instructor's syllabus):

Language Awareness (Eschholtz)
Thomson Reader (Yagelski)
Norton Reader (SHORTER) (Peterson)
Writing in the Works (Burak/Blau)
Presence of Others (Ruszkiewicz)

Students are also encouraged to purchase a good dictionary, such as *Webster's New World College Dictionary* (4th edition).

NOTE: With prior approval of t

he Composition Program Committee, instructors may use an alternative textbook.

III.Course Goals

Upon successful completion of Writing 1, students should be able to:

A. 1. Comprehend, discuss, and critically analyze assigned readings.

B. 2. Communicate effectively in writing by

- responding successfully to a variety of assignments
- using appropriate rhetorical strategies for developing and organizing ideas
- incorporating both instructor and peer feedback in the revision processes
- completing written work that responds appropriately to the assignment and displays a minimum of errors

C. 3. Accessing and use a variety of learning tools and technologies, such as:

- articles, books and other materials at Maag library
- email, the Internet, and word processors (e.g., Microsoft Word)
- academic support services such as the Writing Center

D. Writing 1 should prepare students for Writing 2 (English 1551). In addition to Goals 1-3 above, students should be introduced to processes of research and source documentation styles (e.g., MLA, APA).

E. All assignments should aim to help students respond to texts critically and to write college-level prose, but instructors may use other kinds of readings and writing activities to help students meet this goal.Engagement in the Learning Process

Engagement in the learning process is a key goal of a college education for all students. Instructors are expected to create conditions in Writing 1 that invite active student engage-ment. Students in Writing 1 are expected to:

- **Read all text selections** assigned by the instructor.
- **Do at least three in-depth writing assignments**, each comprising at least one rough draft as well as a final draft that shows evidence of significant revision. The final drafts of these assignments should be approximately 1000-1200 words long (4-6 double-spaced pages).

  Typically, these writing assignments should involve the completion of a formal work in writing that both expresses the student's perspective and shows engagement with some set of readings. However, instructors may use their discretion to customize these assignments, within the parameters of Writing 1's general goals.

  Complete all other assignments required by the in-structor (e.g., in-class or out-of-class writing, graded or ungraded writing, prewriting, and revision or rewriting.)
- **Follow the schedule and policies** in the instructor's syllabus regarding preparation, attendance, classroom par-ticipation, and assignment deadlines.

F. Information Literacy

Information literacy is essential for all YSU students. To en-hance students' information literacy, instructors are expected to introduce students in Writing 1 to resources at Maag Li-brary during at least one class session. Instructors may con-duct these introductions themselves or with the assistance of library staff, who can offer directed, customized instruction.

G. Computer Lab Classrooms

When class is held in a computer lab, instructors are expected to take advantage of technologies useful in engaging students in the activities of the course. Instructors are also expected to introduce students briefly to the hardware and software necessary for completing the course successfully; students requiring additional assistance may be referred to the Center for Student Progress (CSP)/Disability Services office (330-941-1372). At minimum, students are expected to learn to use email, the Internet, and Microsoft Word.

IV. Grades

Final grades for Writing 1 are A, B, C, and NC (No Credit): Students who earn a grade of A or B have fulfilled course assignments at an outstanding or higher than average level; a C indicates satisfactory performance in the course.

A final average of less than C will earn a grade of NC for the course. An NC does not affect the overall GPA, but it does appear on the student's transcript, and the student must repeat Writing 1. It may be retaken only once without the approval of the dean. Successful completion of Writing 1 is required before a student will be permitted to register for Writing 2 (English 1551).

Students are responsible for being aware of their grade in Writing 1 when they register for their next semester's classes.

YSU Policy on Incomplete grades: The instructor may assign a grade of Incomplete (I) only if the following con-ditions are met:

- the student has requested the Incomplete ahead of time;
- all course work prior to this request has been satisfactorily completed;
- the Instructor agrees that an I is warranted.

YSU policy states that students have at maximum one year to complete an Incomplete; instructors are permitted to require that the work be completed in a shorter amount of time. If no formal grade change occurs within one year, the I automatic-ally reverts to an NC.

Audit Policy: Students who register to audit a composition course should consult the instructor about minimum require-ments.

Transfer credit: Students wishing to receive transfer credit for a college composition course taken elsewhere must have that course evaluated by the Composition Program Direc-tor(s) before taking any of YSU's composition courses.

V. Plagiarism and Academic Dishonesty

All assignments completed in English 1550 must be the pro-duct of the student's own thought and inquiry.

Plagiarism means presenting words, ideas, or information found in works written by others as if they were your own.

Academic dishonesty includes plagiarism, all forms of cheating, as well as receiving inappropriate assistance from others in completing an assignment.

Instructors are responsible for teaching students about plagiarism and for distinguishing between intentional plagiar-ism and unintentional errors of citation. However, students are responsible for any actions that might constitute plagiarism. The penalty for plagiarism ranges from failing the assignment to getting an NC in the course. For more information, please see the YSU Student code, particularly Articles I and IX, at http://www.ysu.edu/thecode.pdf.

VI. Academic Support Resources

A. The Writing Center

Students may seek additional one-on-one assistance at the Writing Center by calling (330) 941-3055 to make an appointment with a Writing Center Consultant. The Writing Center has hours between Monday and Friday whenever classes are in session. The Writing Center's website is at:

http://www.as.ysu.edu/~english/wc.html

B. Americans With Disabilities Act

Anyone requiring special adaptations or accommodations should inform the instructor as soon as possible. In accordance with University procedures, if you have a documented disability and require accommodations to obtain equal access in this course, please contact the Office of Disability Services (phone: 330-941-1372) in the Center for Student Progress at the beginning of the semester or when given an assignment for which an accommodation is required. Students with disabilities must verify their eligibility through the Office of Disability Services.

Appendix E
Student Consent Form

Informed Consent Form

Dear Student,

We are conducting a study to compare differences in quality of writing and cognitive levels in papers from English 1550 at Youngstown State University.

In this study, you will be asked permission to give an extra copy of your sample/diagnostic essay and final essay to another professor for data analysis.

There are no risks to you personally.

All data will be handled anonymously, so that no one will be able to identify you personally when the results are recorded.

Your participation is totally voluntary, and you may withdraw at any time.
Should you decide to withdraw, tell your professor.

For further information or to ask questions, please feel free to contact:

Grants & Sponsored Projects	Professor Suzanne Penner
357 Tod Hall	Adjunct English professor
X 2377	233 DeBartolo
	ispenner@ysu.edu

I understand the study described above. I am 18 years of age or older, and I agree to participate.

_____

Signature Date

Appendix F
Professor Permission Form

Participation Form

I understand that I will ask my students to allow another professor to use their essays to conduct a research project.

> I will collect the Sample (diagnostic) essay and the Final essay of the semester from my students for the research project.

I understand that I will offer my students an Informed Consent Form that legally allows another professor to use their essays.

_____

Signature Date

Please mail to:
Suzanne Penner
838 Elmwood Drive
Hubbard, OH 44425

Appendix G
Permission from Dr. Terri Flateby

Note from researcher: Researcher did not need permission to use the CLAQWA instrument because it is available online to anyone, but permission was sought and granted by Dr. Flateby.

RE: Using the CLAQWA - Inbox - Yahoo! Mail Page 1 of 1

Suzanne,

You have my permission. Actually you have a "paper" trail with CLAQWA Online; if your experimental group is using peer review (as most compositions do), we have evidence to suggest its benefit for improving writing. Also, we have added more critical thinking traits to the rubric.

Just some thoughts-

When does your data collection process begin?

Terri Flateby

Terri L. Flateby, Ph.D.

Director, Office of Assessment

University of South Florida

Tampa, FL 33620

813.974.5298 http://M'yw,usf,edu/Assessment

Appendix H
Permission from Youngstown State University

Youngstown
STATE UNIVERSITY
Youngstown, Ohio 44555

One University Plaza,

School of Graduate Studies and Research
Office of the Dean
330.941.3091
Fax 330.941.1580
graduateschool@cc.ysu.edu

January 20, 2009

Ms. Suzanne Penner, Principal Investigator
Department of English
UNIVERSITY

RE: HSRC PROTOCOL NUMBER: 69-2009
TITLE: Differences in Cognitive Level and Quality of Writing in Essays of College Freshmen

Dear Ms. Penner:

The Human Subjects Research Committee of Youngstown State University has reviewed your response to their concerns regarding the above mentioned protocol and determined that your protocol now meets YSU Human Subjects Research guidelines. Therefore, I am pleased to inform you that your project has been fully approved.

Please note that your project is approved for one year. If your project extends beyond one year, you must submit a project Update form at that time.

Any changes in your research activity should be promptly reported to the Human Subjects Research Committee and may not be initiated without HSRC approval except where necessary to eliminate hazard to human subjects. Any unanticipated problems involving risks to subjects should also be promptly reported to the Human Subjects Research Committee.

We wish you well in your study.

Sincerely,

Peter J. Kasvinsky
Associate Provost for Research
Research Compliance (

PJK:cc
Youngstown
y Salvner, Chair
of English
epartment
1908-2008
A PROUD PAST
A PROMISING FUTURE

www.ysu.edu |T-»j

Appendix I
Permission from Liberty University

LIBERTY
UNIVERSITY ™

The Graduate School at Liberty University

IRB Approval 640.100308: Suzanne Penner

Differences in Cognitive Level and Quality of Writing in Research Papers of College Freshmen

March 25, 2010

Dear Suzanne,

We are pleased to inform you that your above study has been approved by the Liberty IRB. This approval is extended to you for one year. If data collection proceeds past one year, or if you make changes in the methodology as it pertains to human subjects, you must resubmit the study to the IRB. See the IRB website for appropriate forms in these cases.

Thank you for your cooperation with the IRB and we wish you well with your research project.

Sincerely,

Fernando Garzon, Psy.D.
[RB Chair, Liberty University
Center for Counseling and Family Studies
Liberty University
1971 University Boulevard
Lynchburg, VA 24502-2269
(434) 592-4054
Fax: (434) 522-0477

address *797?* University Boulevard Lynchburg, VA 24502
phone fax
434-592-4044 434-522-0506
web www.liberty.edu/academics/graduate

Table 1
Raw Data - SPSS Variables List

```
Data written to C:\Documents and
Settings\...\Dissertation\DissExcel.xls.
14 variables and 107 cases written to range: SPSS.

Variable: id                Type: Number   Width:  1   Dec: 0
Variable: lastname          Type: String   Width: 10
Variable: firstnam          Type: String   Width: 10
Variable: sex               Type: Number   Width:  1   Dec: 0
Variable: diag1             Type: Number   Width:  8   Dec: 2
Variable: diag2             Type: Number   Width:  8   Dec: 2
Variable: final1            Type: Number   Width:  8   Dec: 2
Variable: final2            Type: Number   Width:  8   Dec: 2
Variable: wordsk1           Type: Number   Width:  8   Dec: 2
Variable: wordsk2           Type: Number   Width:  8   Dec: 2
Variable: reason1           Type: Number   Width:  8   Dec: 2
Variable: reason2           Type: Number   Width:  8   Dec: 2
```

Table 2
Raw Data - SPSS Data File

Numeric ID of student	Gender 1=female 2=male	Diag 1 Score	Diag 2 Score	Final 1 Score	Final 2 Score	Word Skill 1	Word Skill 2	Reason 1	Reason 2
101	1	15.00	13.00	35.00	28.00	12.00	16.00	15.00	20.00
102	1	14.00	13.00	35.00	27.00	12.00	15.00	15.00	20.00
103	1	15.00	18.00	26.00	27.00	11.00	16.00	11.00	15.00
104	1	20.00	12.00	35.00	20.00	8.00	12.00	15.00	20.00
105	2	25.00	20.00	35.00	24.00	10.00	14.00	15.00	20.00
106	1	17.00	15.00	35.00	33.00	13.00	20.00	15.00	20.00
107	2	17.00	10.00	26.00	23.00	10.00	13.00	10.00	16.00
108	1	17.00	18.00	28.00	23.00	11.00	12.00	12.00	16.00
109	2	12.00	13.00	28.00	26.00	11.00	15.00	12.00	16.00
110	1	16.00	15.00	35.00	25.00	11.00	14.00	15.00	20.00
111	1	17.00	14.00	35.00	24.00	11.00	13.00	15.00	20.00
112	2	12.00	11.00	14.00	21.00	5.00	16.00	6.00	8.00
113	2	20.00	16.00	21.00	20.00	8.00	12.00	9.00	12.00
114	1	21.00	10.00	21.00	28.00	12.00	16.00	9.00	12.00
115	2	24.00	19.00	22.00	19.00	7.00	12.00	9.00	13.00
116	1	20.00	16.00	14.00	27.00	11.00	16.00	6.00	8.00
117	2	20.00	17.00	14.00	28.00	12.00	16.00	6.00	8.00
118	1	24.00	16.00	21.00	19.00	6.00	13.00	9.00	12.00
119	2	19.00	12.00	15.00	19.00	8.00	11.00	7.00	8.00
120	1	20.00	16.00	21.00	27.00	11.00	16.00	9.00	12.00
121	2	19.00	15.00	21.00	26.00	11.00	15.00	9.00	12.00
122	2	18.00	14.00	20.00	24.00	10.00	14.00	8.00	12.00
123	2	14.00	13.00	35.00	25.00	11.00	14.00	15.00	20.00
124	1	10.00	13.00	28.00	24.00	8.00	16.00	12.00	16.00
125	2	21.00	16.00	21.00	26.00	13.00	13.00	9.00	12.00
126	2	19.00	14.00	35.00	24.00	9.00	13.00	15.00	20.00
127	1	14.00	11.00	28.00	27.00	12.00	15.00	12.00	16.00
128	2	22.00	14.00	25.00	26.00	12.00	14.00	10.00	15.00
129	2	12.00	8.00	29.00	23.00	11.00	12.00	13.00	16.00
131	2	9.00	11.00	35.00	25.00	11.00	14.00	15.00	20.00
132	1	17.00	15.00	23.00	20.00	8.00	12.00	9.00	14.00
133	1	25.00	19.00	35.00	28.00	12.00	16.00	15.00	20.00
134	2	16.00	15.00	23.00	15.00	6.00	9.00	10.00	13.00

135	1	14.00	12.00	25.00	25.00	11.00	14.00	12.00	13.00
136	2	18.00	12.00	35.00	29.00	13.00	16.00	15.00	20.00
137	1	10.00	10.00	20.00	16.00	7.00	9.00	8.00	12.00
138	1	14.00	15.00	28.00	30.00	14.00	16.00	12.00	16.00
139	2	13.00	10.00	35.00	25.00	11.00	14.00	15.00	20.00
140	2	14.00	14.00	21.00	25.00	11.00	14.00	9.00	12.00
141	2	17.00	11.00	35.00	23.00	10.00	13.00	15.00	20.00
142	2	13.00	14.00	21.00	18.00	8.00	10.00	9.00	12.00
143	2	21.00	20.00	28.00	23.00	11.00	12.00	12.00	16.00
145	1	14.00	13.00	28.00	28.00	12.00	16.00	12.00	16.00
146	1	19.00	17.00	21.00	24.00	10.00	14.00	9.00	12.00
147	2	13.00	13.00	21.00	23.00	9.00	14.00	9.00	12.00
148	2	15.00	17.00	28.00	27.00	12.00	15.00	12.00	16.00
149	2	25.00	18.00	28.00	24.00	10.00	14.00	12.00	16.00
150	1	25.00	18.00	28.00	27.00	12.00	15.00	12.00	16.00
151	1	12.00	13.00	26.00	28.00	12.00	16.00	11.00	15.00
152	2	20.00	20.00	21.00	30.00	14.00	16.00	9.00	12.00
153	1	18.00	20.00	28.00	28.00	12.00	16.00	12.00	16.00
154	1	20.00	13.00	21.00	29.00	12.00	17.00	9.00	12.00
155	1	6.00	8.00	21.00	21.00	8.00	13.00	9.00	12.00
156	1	25.00	13.00	27.00	27.00	12.00	15.00	12.00	15.00
157	1	9.00	10.00	14.00	24.00	10.00	14.00	6.00	8.00
201	1	14.00	10.00	30.00	23.00	9.00	14.00	14.00	16.00
202	1	12.00	10.00	26.00	21.00	7.00	14.00	12.00	14.00
203	2	11.00	10.00	28.00	27.00	12.00	15.00	12.00	16.00
204	2	17.00	14.00	28.00	28.00	12.00	16.00	12.00	16.00
205	1	17.00	14.00	33.00	28.00	12.00	16.00	14.00	19.00
206	2	14.00	12.00	28.00	26.00	10.00	14.00	12.00	16.00
207	1	25.00	15.00	28.00	21.00	7.00	13.00	12.00	16.00
208	1	16.00	13.00	21.00	25.00	11.00	14.00	9.00	12.00
209	2	22.00	15.00	21.00	24.00	10.00	14.00	9.00	12.00
210	1	11.00	11.00	21.00	27.00	12.00	15.00	9.00	12.00
211	1	12.00	10.00	29.00	24.00	10.00	14.00	12.00	17.00
212	2	14.00	13.00	14.00	19.00	8.00	11.00	6.00	8.00
213	1	5.00	8.00	21.00	22.00	10.00	12.00	9.00	12.00
214	2	7.00	10.00	21.00	28.00	12.00	16.00	9.00	12.00
215	1	25.00	16.00	30.00	24.00	11.00	13.00	13.00	17.00
216	1	12.00	18.00	30.00	28.00	12.00	16.00	13.00	17.00
217	2	16.00	17.00	28.00	23.00	11.00	12.00	12.00	16.00
218	2	18.00	19.00	24.00	26.00	11.00	15.00	10.00	14.00
219	2	15.00	17.00	14.00	25.00	10.00	15.00	6.00	8.00
220	2	19.00	21.00	21.00	26.00	12.00	14.00	9.00	12.00
221	2	14.00	15.00	21.00	22.00	9.00	13.00	9.00	12.00
222	2	13.00	12.00	21.00	27.00	12.00	15.00	9.00	12.00
223	2	15.00	15.00	21.00	21.00	9.00	12.00	9.00	12.00
224	1	18.00	18.00	28.00	26.00	11.00	15.00	12.00	16.00

225	2	16.00	15.00	26.00	22.00	8.00	14.00	10.00	16.00
226	2	20.00	19.00	15.00	27.00	12.00	15.00	7.00	8.00
227	2	14.00	13.00	28.00	26.00	10.00	16.00	12.00	16.00
228	1	17.00	15.00	26.00	26.00	12.00	14.00	10.00	16.00
229	2	19.00	17.00	28.00	24.00	12.00	12.00	12.00	16.00
230	2	19.00	12.00	14.00	23.00	9.00	14.00	6.00	8.00
231	1	11.00	13.00	21.00	20.00	9.00	11.00	9.00	12.00
232	2	20.00	19.00	29.00	27.00	12.00	15.00	13.00	16.00
233	2	10.00	13.00	19.00	24.00	10.00	14.00	9.00	10.00
234	1	24.00	16.00	28.00	26.00	12.00	14.00	12.00	16.00
235	2	24.00	22.00	28.00	24.00	10.00	14.00	12.00	16.00
236	2	5.00	10.00	7.00	26.00	12.00	16.00	3.00	4.00
237	2	25.00	22.00	21.00	28.00	12.00	16.00	9.00	12.00
238	1	22.00	25.00	28.00	26.00	12.00	14.00	12.00	16.00
239	2	25.00	15.00	28.00	28.00	12.00	16.00	12.00	16.00
240	2	21.00	19.00	21.00	23.00	11.00	12.00	9.00	12.00
241	2	25.00	21.00	28.00	26.00	11.00	15.00	12.00	16.00
242	1	15.00	14.00	21.00	27.00	11.00	16.00	9.00	12.00
243	1	17.00	12.00	35.00	25.00	11.00	14.00	15.00	20.00
244	1	23.00	18.00	30.00	26.00	11.00	15.00	12.00	18.00
245	1	17.00	15.00	35.00	26.00	11.00	15.00	15.00	20.00
246	1	13.00	13.00	35.00	28.00	12.00	16.00	15.00	20.00
247	1	15.00	11.00	21.00	25.00	11.00	14.00	9.00	12.00
248	1	15.00	16.00	35.00	28.00	12.00	16.00	15.00	20.00
249	1	12.00	12.00	21.00	27.00	12.00	15.00	9.00	12.00
250	1	23.00	13.00	21.00	20.00	6.00	14.00	9.00	12.00
251	2	24.00	13.00	35.00	23.00	12.00	11.00	15.00	20.00
252	2	22.00	12.00	28.00	25.00	11.00	14.00	12.00	16.00

Table 3
Summary of Results - Statistical Analysis from Steven McDonald, M.B.A.
The tables and analyses on the following pages were completed by the statistician.

Table 3.1
Inter rater reliability for Diagnostic Essay

Descriptive Statistics

	Mean	Std. Deviation	N
diag1	16.9252	4.90610	107
diag2	14.4953	3.39637	107

Correlations

		diag1	diag2
diag1	Pearson Correlation	1	.636**
	Sig. (2-tailed)		.000
	N	107	107
diag2	Pearson Correlation	.636**	1
	Sig. (2-tailed)	.000	
	N	107	107

**. Correlation is significant at the 0.01 level (2-tailed).

Both graders showed a statistically significant correlation on the grading of the diagnostic essay. The Pearson correlation was .636 (sig. = 0.00) at Ŭ= 0.01 in a two-tailed test.

Table 3.2
Inter Rater Reliability for Final Essay

Descriptive Statistics

	Mean	Std. Deviation	N
final1	25.5421	6.33541	107
final2	24.7850	3.09604	107

Correlations

		final1	final2
final1	Pearson Correlation	1	.252**
	Sig. (2-tailed)		.009
	N	107	107
final2	Pearson Correlation	.252**	1
	Sig. (2-tailed)	.009	
	N	107	107

**. Correlation is significant at the 0.01 level (2-tailed).

Both graders showed a statistically significant correlation on the grading of the final essay. The Pearson correlation was .252 (sig. = 0.09) at Ŭ= 0.01 in a two-tailed test.

Table 3.3
One Way ANOVA for Diagnostic Essay (experimental and control groups)

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
diag1	experimental	55	17.0182	4.60471	.62090	15.7734	18.2630	6.00	25.00
	control	52	16.8269	5.24954	.72798	15.3654	18.2884	5.00	25.00
	Total	107	16.9252	4.90610	.47429	15.9849	17.8656	5.00	25.00
diag2	experimental	55	14.2364	3.10891	.41921	13.3959	15.0768	8.00	20.00
	control	52	14.7692	3.68684	.51127	13.7428	15.7957	8.00	25.00
	Total	107	14.4953	3.39637	.32834	13.8444	15.1463	8.00	25.00

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
diag1	.900	1	105	.345
diag2	1.077	1	105	.302

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
diag1	Between Groups	.978	1	.978	.040	.841
	Within Groups	2550.424	105	24.290		
	Total	2551.402	106			
diag2	Between Groups	7.590	1	7.590	.656	.420
	Within Groups	1215.158	105	11.573		
	Total	1222.748	106			

ANOVA is robust to heterogeneity of variance if sample sizes are equal or near equal. However, a Levene analysis was run to confirm homogeneity of variance in the samples. The

Levene statistic tests the null hypothesis of homogeneity of variance with Ŭ< .10 being needed to reject the null hypothesis. Both samples failed to reject the null hypothesis of homogeneity of variance.

There were no statistically significant differences between the experimental and control groups for the diagnostic essays graded by rater one or rater two. With an alpha level of Ŭ= 0.05, diagnostic essay one measured Ŭ= 0.841 and diagnostic essay two measured Ŭ= 0.420. Thus, the null hypothesis of no significant difference between the experimental and control groups in performance on the diagnostic essay was not rejected.

Table 3.4
Word Skills Analysis for Final Essay (experimental and control groups)

The study in question sought to test for statistically significant differences between the experimental group (N = 55) and the control group (N = 52) on the word skills portion of the Cognitive Level and Quality Writing Assessment system (CLAQWA) instrument. The analysis was further discriminated by the use of two independent graders.

Word Skills Analysis

Descriptives

		N	Mean	Std. Deviation	Std. Error
wordsk1	experimental	55	10.4545	2.03505	.27441
	control	52	10.6923	1.52802	.21190
	Total	107	10.5701	1.80205	.17421
wordsk2	experimental	55	14.1455	2.02227	.27268
	control	52	14.2308	1.42272	.19730
	Total	107	14.1869	1.74902	.16908

The descriptive statistics show little significant difference between the mean scores for the experimental and control groups as reflected between the two groups and the overall mean score. This pattern holds for both graders.

The Levene statistic was used to test for homogeneity of variance. The null hypothesis of the test of homogeneity of variance is that there is no significant difference between the variances of the two groups. The hypothesis is tested at Ŭ= .10.

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
wordsk1	3.855	1	105	.052
wordsk2	4.244	1	105	.042

The null hypothesis of no difference between the variances of the experimental and control groups was rejected for both graders in the word skills portion of the test. However, this does not prove to be problematic for an ANOVA analysis since the sample sizes are almost equal. ANOVA is robust to heterogeneity of variance if the groups sizes are equal or near to equal.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
wordsk1	Between Groups	1.511	1	1.511	.463	.498
	Within Groups	342.713	105	3.264		
	Total	344.224	106			
wordsk2	Between Groups	.195	1	.195	.063	.802
	Within Groups	324.067	105	3.086		
	Total	324.262	106			

The ANOVA shows no statistically significant difference between the two groups measured as wordsk1 (Grader 1). The alpha level for the analysis was Ŭ= .05. The between groups significance was Ŭ= .498, thus failing to reject the null hypothesis of no significant difference between the two groups. This also held true for the groups designated wordsk2 (Grader 2). The between groups significance was Ŭ= .802, thus failing to reject the null hypothesis of no significant difference between the groups.

Table 3.5
Reasoning Skills Analysis for Final Essay (experimental and control groups)

Reasoning Skills Analysis

Descriptives

		N	Mean	Std. Deviation	Std. Error
reason1	experimental	55	11.1455	2.84422	.38351
	control	52	10.7115	2.61471	.36260
	Total	107	10.9346	2.73092	.26401
reason2	experimental	55	14.9273	3.74094	.50443
	control	52	14.2692	3.54310	.49134
	Total	107	14.6075	3.64396	.35227

The descriptive statistics show little significant difference between the mean scores for the experimental and control groups as reflected between the two groups and the overall mean score. This pattern holds for both graders.

The Levene statistic was used to test for homogeneity of variance. The null hypothesis of the test of homogeneity of variance is that there is no significant difference between the variances of the two groups. The hypothesis is tested at Ŭ= .10.

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
reason1	.826	1	105	.366
reason2	.171	1	105	.680

The null hypothesis of homogeneity of variance was not rejected. It can be safely assumed that the two groups, for each grader, possessed equal or near equal variances.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
reason1	Between Groups	5.033	1	5.033	.673	.414
	Within Groups	785.509	105	7.481		
	Total	790.542	106			
reason2	Between Groups	11.574	1	11.574	.871	.353
	Within Groups	1395.940	105	13.295		
	Total	1407.514	106			

The ANOVA shows no statistically significant difference between the two groups measured as reason1 (Grader 1). The alpha level for the analysis was Ŭ= .05. The between groups significance was Ŭ= .414, thus failing to reject the null hypothesis of no significant difference between the two groups. This also held true for the groups designated reason2 (Grader 2). The between groups significance was Ŭ= .353, thus failing to reject the null hypothesis of no significant difference between the groups.

Summary

Based on the results of these statistics it can be concluded that no statistically significant difference exists between the experimental group and the control group used in the analysis. This held true for the evaluations given by both Grader 1 and Grader 2.

Table 3.6
One Way ANOVA for Final Essay (experimental and control groups)

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
final1	experimental	55	26.0727	6.55138	.88339	24.3016	27.8438	14.00	35.00
	control	52	24.9808	6.11167	.84754	23.2793	26.6823	7.00	35.00
	Total	107	25.5421	6.33541	.61247	24.3278	26.7563	7.00	35.00
final2	experimental	55	24.6364	3.63809	.49056	23.6529	25.6199	15.00	33.00
	control	52	24.9423	2.42061	.33568	24.2684	25.6162	19.00	28.00
	Total	107	24.7850	3.09604	.29931	24.1916	25.3784	15.00	33.00

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
final1	.351	1	105	.555
final2	5.407	1	105	.022

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
final1	Between Groups	31.871	1	31.871	.792	.375
	Within Groups	4222.690	105	40.216		
	Total	4254.561	106			
final2	Between Groups	2.502	1	2.502	.259	.612
	Within Groups	1013.554	105	9.653		
	Total	1016.056	106			

ANOVA is robust to heterogeneity of variance if sample sizes are equal or near equal. However, a Levene analysis was run to confirm homogeneity of variance in the samples. The

Levene statistic tests the null hypothesis of homogeneity of variance with Ŭ< .10 being needed to reject the null hypothesis. Both samples failed to reject the null hypothesis of homogeneity of variance.

There were no statistically significant differences between the experimental and control groups for the final essays graded by rater one or rater two. With an alpha level of Ŭ= 0.05, final essay one measured Ŭ= 0.375 and final essay two measured Ŭ= 0.612. Thus, the null hypothesis of no significant difference between the experimental and control groups in performance on the final essay was not rejected.

Table 3.7
One Way ANOVA for Diagnostic Essay (female and male)

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
diag1	female	52	16.5192	5.00388	.69391	15.1261	17.9123	5.00	25.00
	male	55	17.3091	4.82614	.65076	16.0044	18.6138	5.00	25.00
	Total	107	16.9252	4.90610	.47429	15.9849	17.8656	5.00	25.00
diag2	female	52	14.0769	3.25303	.45111	13.1713	14.9826	8.00	25.00
	male	55	14.8909	3.51016	.47331	13.9420	15.8398	8.00	22.00
	Total	107	14.4953	3.39637	.32834	13.8444	15.1463	8.00	25.00

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
diag1	.000	1	105	.993
diag2	.708	1	105	.402

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
diag1	Between Groups	16.676	1	16.676	.691	.408
	Within Groups	2534.726	105	24.140		
	Total	2551.402	106			
diag2	Between Groups	17.710	1	17.710	1.543	.217
	Within Groups	1205.038	105	11.477		
	Total	1222.748	106			

ANOVA is robust to heterogeneity of variance if sample sizes are equal or near equal. However, a Levene analysis was run to confirm homogeneity of variance in the samples. The

Levene statistic tests the null hypothesis of homogeneity of variance with Ŭ< .10 being needed to reject the null hypothesis. Both samples failed to reject the null hypothesis of homogeneity of variance.

There were no statistically significant differences between the female and male groups for the diagnostic essays graded by rater one or rater two. With an alpha level of Ŭ= 0.05, diagnostic essay one measured Ŭ= 0.408 and diagnostic essay two measured Ŭ= 0.217. Thus, the null hypothesis of no significant difference between the male and female groups in performance on the diagnostic essay was not rejected.

Table 3.8
One Way ANOVA for Final Essay (female and male)

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
final1	female	52	26.8462	5.78160	.80176	25.2365	28.4558	14.00	35.00
	male	55	24.3091	6.63569	.89476	22.5152	26.1030	7.00	35.00
	Total	107	25.5421	6.33541	.61247	24.3278	26.7563	7.00	35.00
final2	female	52	25.2500	3.21074	.44525	24.3561	26.1439	16.00	33.00
	male	55	24.3455	2.94529	.39714	23.5492	25.1417	15.00	30.00
	Total	107	24.7850	3.09604	.29931	24.1916	25.3784	15.00	33.00

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
final1	1.187	1	105	.278
final2	.510	1	105	.477

ANOVA

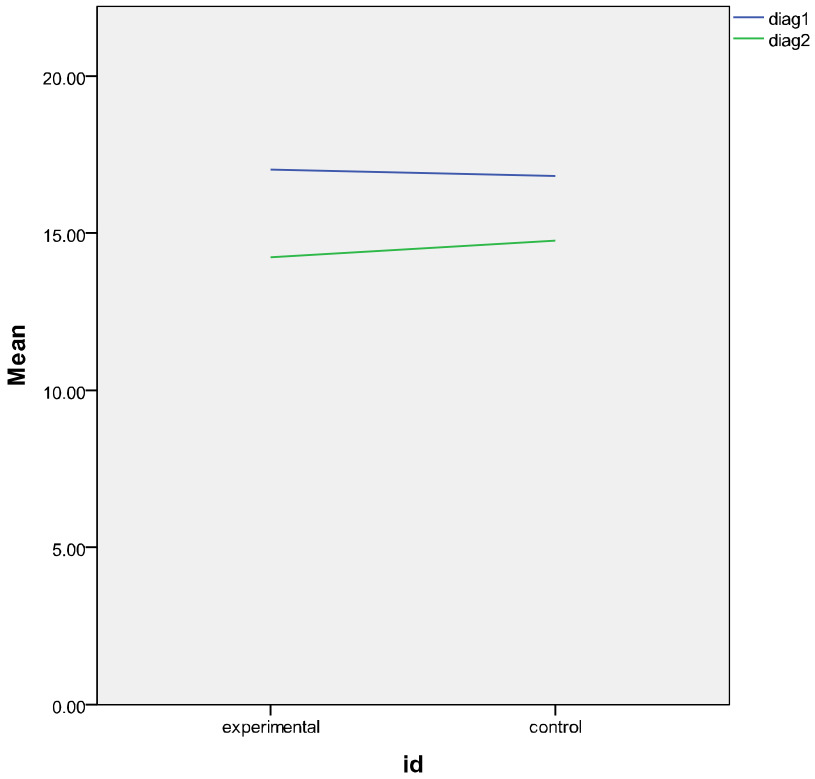
		Sum of Squares	df	Mean Square	F	Sig.
final1	Between Groups	172.046	1	172.046	4.425	.038
	Within Groups	4082.515	105	38.881		
	Total	4254.561	106			
final2	Between Groups	21.870	1	21.870	2.310	.132
	Within Groups	994.186	105	9.468		
	Total	1016.056	106			

ANOVA is robust to heterogeneity of variance if sample sizes are equal or near equal. However, a Levene analysis was run to confirm homogeneity of variance in the samples. The Levene statistic tests the null hypothesis of homogeneity of variance with Ŭ< .10 being needed to reject the null hypothesis. Both samples failed to reject the null hypothesis of homogeneity of variance.

There was a statistically significant difference between the female and male groups for the diagnostic essays graded by rater one. With an alpha level of Ŭ= 0.05, diagnostic essay one measured Ŭ= 0.038. However, there was not a statistically significant difference between male and female groups on the final essay as graded by rater two. With an alpha level of Ŭ= 0.05 diagnostic essay two measured Ŭ= 0.132.

While this proves to be interesting, and may warrant further analysis, it may rove to be problematic in rejecting the null hypothesis of no significant difference between male and female groups on the final essay. This is especially true in light of the results of inter rater reliability for both graders. Thus, the null hypothesis of no significant difference between the male and female groups in performance on the diagnostic essay was not rejected.

Figure 1
Graphs of Scores for Control and Experimental Groups



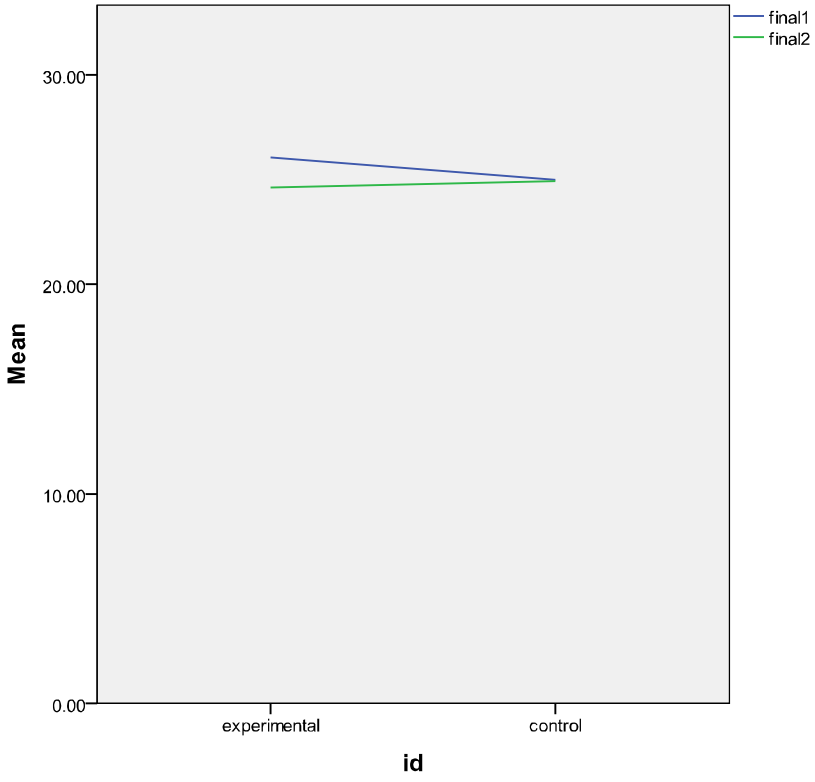


Figure 2
Graphs of Scores for Female and Male Groups