



1985

An Empirical Study of Student Compositions and Their Improvement

Donna Van de Water
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_theses



Part of the [Psychology Commons](#)

Recommended Citation

Van de Water, Donna, "An Empirical Study of Student Compositions and Their Improvement" (1985).
Master's Theses. 3392.
https://ecommons.luc.edu/luc_theses/3392

This Thesis is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Master's Theses by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).
Copyright © 1985 Donna Van de Water

AN EMPIRICAL STUDY OF STUDENT COMPOSITIONS
AND THEIR IMPROVEMENT

by

Donna Van De Water

A Thesis Submitted to the Faculty of the Graduate School
of Loyola University of Chicago in Partial Fulfillment
of the Requirements for the Degree of
Master of Arts

May

1985

ACKNOWLEDGMENTS

The outstanding cooperation, encouragement and guidance from my committee members Daniel C. O'Connell, S.J. and Deborah L. Holmes is gratefully acknowledged. Thanks and appreciation is extended to Dan O'Connell for his constant support and for allowing me to take over this research. A special note of thanks is extended to Mr. and Mrs. Duerr. They will never know how much I appreciated them allowing me to sit at their word processor for days on end to get this done. Additional thanks is extended to Todd Miller and Leslie Scott for helping me to score essays.

Finally, I would like to dedicate this thesis to Jim. Without his unconditional love and support I doubt if I would have had the courage to complete such a project.

VITA

The author, Donna A. Van De Water, is the daughter of Donald O. Van De Water and Phyllis J. Van De Water. She was born on May 1, 1959.

Her high school diploma was received from Watchung Hills Regional High School in Warren, New Jersey in 1977. In May, 1981 she received the degree of Bachelor of Arts with majors in Sociology and Psychology from Valparaiso University in Indiana.

From June, 1981 until August, 1983 Ms. Van De Water worked for Drexel Burnham Lambert in Chicago, Illinois. She was employed as a sales assistant and was promoted to Office Administrator. In August of 1983 she terminated her employment so as to attend Loyola University of Chicago full time. Since that time she has worked as a research assistant in the Psychology Department at Loyola University.

TABLE OF CONTENTS

| | Page |
|--|------|
| ACKNOWLEDGMENTS | ii |
| VITA | iii |
| LIST OF TABLES | vi |
| LIST OF FIGURES | vii |
| Chapter | |
| I. INTRODUCTION | 1 |
| Writing Across the Curriculum and Improving Student Writing | 1 |
| Statement of the Problem | 2 |
| Need for the Study | 4 |
| Description of the Study | 5 |
| Limitations of the Study | 5 |
| Overview of the Thesis | 6 |
| II. REVIEW OF THE LITERATURE | 8 |
| Definition of Writing | 8 |
| Composition Processes | 11 |
| Revision | 23 |
| Improving Student Writing | 26 |
| Methods of Writing Evaluation | 31 |
| Frequency of Micro-elements and Essay Writing | 34 |
| Summary | 35 |
| III. METHODOLOGY | 38 |
| Subjects | 38 |
| Materials | 39 |
| Design and Procedure | 40 |
| Preparation of Essays for Analysis | 41 |
| The Reading and Scoring Session | 42 |
| Dependent Measures | 44 |
| Questions To Be Answered | 44 |
| Data Analysis | 50 |

| | Page |
|---|------|
| IV. RESULTS AND DISCUSSION | 52 |
| Descriptive Summary Statistics | 52 |
| Intrareader Reliability | 60 |
| Interreader Reliability | 62 |
| Differences Between Control and Experimental Groups . . | 65 |
| Differences Between First and Revised Essays | 66 |
| Concurrent Validity and Construct Validity | 67 |
| Summary of Findings | 68 |
| V. SUMMARY, CONCLUSIONS, AND IMPLICATIONS | 71 |
| Summary of Design | 71 |
| Findings | 73 |
| Conclusions | 76 |
| Implications for Education | 79 |
| Implications for Research | 80 |
| REFERENCES | 82 |
| APPENDIX A | 87 |
| Sample Essays | 88 |
| Sample Score Card | 92 |
| APPENDIX B | 93 |
| Raw Data | 94 |

LIST OF TABLES

| Table | Page |
|--|------|
| 1. Mean and Standard Deviation of Reader 1 and Reader 2 Assigned Scores and Total Scores for Control (<u>N</u> =16) and Experimental Groups (<u>N</u> =15) on the First and Revised Essays | 56 |
| 2. Mean and Standard Deviations for Number of Words, Syllables per Word and Token-Type Ratio for the Control and Experimental Groups Across Essays (First and Revised) | 58 |
| 3. Doubly Multivariate Analysis of Variance Results Between Control and Experimental Groups for Micro-Elements (Number of Words, Syllables per Word, and Token-Type Ratio) on First and Revised Essays | 59 |
| 4. Intrareader Reliability Coefficients for Reader 1 and Reader 2 by Control (<u>N</u> =16) and Experimental (<u>N</u> =15) Groups Across First and Revised Essays | 61 |
| 5. Discrepant and Nondiscrepant Assigned Scores by Control and Experimental Groups, Across First and Revised Essays, and by Reader 1 and Reader 2 | 63 |
| 6. Interreader Reliability Coefficients of Total Essay Scores for Reader 1 and Reader 2 by Control (<u>N</u> =16) and Experimental (<u>N</u> =15) Groups Across First and Revised Essays | 64 |
| 7. Pearson Product Moment Correlations Between Writing I and General Psychology and Micro-Elements (Number of Words, Syllables per Word, and Token-Type Ratio) of First and Revised Essays | 69 |

LIST OF FIGURES

| Figure | Page |
|---|------|
| 1. Distribution of Total Assigned Scores by Essay (First and Revised) | 53 |
| 2. Distribution of Assigned Scores for First Essays by Reader 1 and Reader 2 | 54 |
| 3. Distribution of Assigned Scores for Revised Essays by Reader 1 and Reader 2 | 55 |

CHAPTER I

INTRODUCTION

Writing Across the Curriculum and Improving Student Writing

There is a consensus that college student writers cannot write as well as they should, and that educators must develop ways to help them improve their writing. It is to this end that the concepts behind writing across the curriculum have been invoked. By incorporating writing instruction in classes other than those in the English department, it is hoped that student writers can show and maintain improvement over time.

The primary goal of writing across the curriculum is to train educators in all disciplines to assist students with their writing. This training would enable educators to feel confident in their text production, and provide correct, constructive criticism. The net result would be improved student writing skills, and maintenance of good writing performance over time.

This study will investigate two procedures advocated by proponents of writing across the curriculum: the use of cues in the writing assignment and revision of essays, in a class outside of the English department. Essays written with and without cues will be compared and revised essays will be compared with their first draft counterparts. This study is also a test of the methodology used in writing assessment. Holistic essay scoring has proven itself

to be reliable and valid in tightly controlled situations; this study will be a test of its abilities in the more true-to-life situation. The essays will also be assessed with quantifiable methods, and these will be compared with the scores produced via holistic essay scoring.

Statement of the Problem

With reports of the National Assessment of Educational Progress-Writing (NAEP-W) and with the extensive media commentary on "Why Johnny Can't Write," the American public has been greatly concerned about students' writing ability. Evidence from NAEP-W and from several other sources has led people to conclude that students are not writing as well as they once did. And a number of experts, legislators, and concerned citizens alike are telling us about the present "crisis" in writing. It does little good to point out that this crisis is not new or to argue that our current understanding of it is derived from questionable data. Nor does it help to tell critics that writing is a complex skill, one that is hard to teach and perhaps even harder to assess adequately. All of these assertions are true. But none of them responds to a growing sense that the public, not to mention our students, deserves a comprehensive assessment of our students' competence as writers (Odell, 1981, p. 95).

The above statement summarized the findings of educators, researchers and parents about the present state of student writing in the United States. Nearly twenty years of declining SAT, ECT, ACT and GRE verbal scores and English composition scores of incoming students at colleges and universities have indicated that perhaps students are not writing as well as their counterparts of the past (Bamberg, 1978; Gelb, 1982; Hendrix, 1981). This apparent decline has resulted in a consequent joining of forces to study writing in the schools.

At the college and university level, the concept of "Writing Across the Curriculum" has become one approach for improvement.

Typically, writing at the college and university level has been delegated to the English department and composition courses are taken during the freshman year. This is referred to as a horizontal approach to writing (Gelb, 1982). The net result of this approach is that writing skills deteriorate from lack of practice over time. A study conducted by Harvard University in 1978 found that seniors majoring in the Natural Sciences were poorer writers than their freshman counterparts; however, seniors majoring in the humanities were better writers than their freshman counterparts, presumably because of practice effects of writing (cited in Gelb, 1982).

The concept of writing across the curriculum emphasizes a vertical sequence of writing instruction. It includes recognition of the following: All instructors of writing, whether specialists or generalists are trained; mature students should be given the opportunity to explain their discipline to general readers in common university discourse; mature students should also be given the opportunity to compose in the language of their discipline for peers and superiors; literacy should be a primary concern of all faculty since it is the basis of higher education; and a system of accountability at all levels of this vertical sequence should be incorporated (Gelb, 1982). Writing across the curriculum involves an understanding that writing is a way of knowing regardless of the discipline (Glatthorn, 1981). Compositions in disciplines other than English are vital; students cannot be permitted to believe that good writing is important only for the English class. Glatthorn (1981) explained that the processes involved in composition require

that writers systemize what they know. This systemization aids in the discovery of new insights. Since writing is an active process, and assuming active processes facilitate learning, it is almost inevitable that writing will result in increased learning.

Need for the Study

During the 1983 fall semester, representatives of each department in the College of Arts and Sciences at Loyola University of Chicago were encouraged to incorporate the concepts learned in a writing across the curriculum workshop into their classrooms. At the end of the semester, they were to report back to the members of the Faculty Advisors to the Core Curriculum Committee. Discussion about the various attempts to incorporate the techniques was very informal with no follow-up.

Daniel C. O'Connell, S.J., representative for the Department of Psychology, expressed dissatisfaction with the methodology used by the members of the Committee of Faculty Advisors for the Core Curriculum. He proposed formalizing procedures so as to assess whether incorporating those concepts recommended by the proponents of writing across the curriculum is indeed effective. The procedures described in this thesis are those used by O'Connell during the Spring semester 1984 at Loyola University of Chicago in a General Psychology course.

Parallel to the need for assessing the particular procedure employed at Loyola, the National Institute of Education (1984) has declared the following research topic to be in need of study:

Research on effective teaching practices for writing, taking into account the content of those practices, the context in which those practices are applied, the effects of those practices when compared with other practices with differing features. All activities should be conducted with the express purpose of presenting the underlying principles which make the practices successful and which can lead to the effective implementation of the practices in other settings (p. 30).

It is hoped that this study will provide some clarifying information deemed necessary by the NIE.

Description of the Study

General Psychology at Loyola University of Chicago is a course primarily taken by students during their freshman year in order to fulfill a general education requirement for graduation. Since there are a large number of students in each class, tests are objective, and little, if any, composition is required.

In addition to regular class testing procedures, students enrolled in General Psychology instructed by Daniel C. O'Connell, S.J. were asked to read a short article and given instructions to write a review or critique. Later in the semester they were asked to revise or rewrite their essays. It was from these students and essays that the data for this thesis were collected.

Limitations of the Study

Writing across the curriculum is based upon the premise that students can become better writers when they practice. This practice is expected to characterize all classes and is to last for the duration of the post-secondary education. The analyses conducted as part of this study are only a small segment of those proposed in the vertical writing sequence suggested by writing across the curriculum.

An effective evaluation of writing across the curriculum would necessitate following students throughout their post-secondary education. Such an evaluation was not possible at the writing of this thesis. It is doubtful if any rigorous study could be conducted on any program utilizing the concepts proposed by writing across the curriculum until formalized and standardized procedures are adopted. Without such standardization, comparisons and evaluations would prove to be invalid and idiosyncratic.

However, by analyzing the results of one experimental procedure it may become possible to develop new methodologies, improve existing methodologies, develop procedures for evaluation, and propose suggestions for further research.

Overview of the Thesis

Chapter I included a statement of the problem, the need for the study, a description of the study, and its limitations.

Chapter II will proceed with a review of the literature. The definition of writing will be explored as well as the processes involved in composition. A special emphasis will be placed on revision. Proposed methods of improving student writing will then be reviewed. A review of existing procedures for assessing writing and the quantifiable features of micro-elements and their relationship to writing assessment will also be included in this section. Maximizing essay scoring reliability and validity will then be discussed. Established methods of measuring growth in writing ability over time will also be included. The chapter will conclude with a brief summary.

Chapter III reviews the methodology of the study. This section

will begin with a description of the subjects, the materials, the design and procedures, the preparation of the essays for analysis, and the reading and scoring session used in the study. A review of the dependent measures will follow. The questions that the thesis hopes to answer as well as hypotheses will be outlined. An explanation of the data analyses of the study will conclude the chapter.

The results of the study will be examined in Chapter IV. The results of the study are based upon descriptive summary statistics, intrareader and interreader reliability, and differences between the control and experimental groups and concurrent and construct validity. The discussion will explore the question as to whether improvement in writing ability over time was facilitated for student writers and the effectiveness of the intervention procedure. Chapter IV will conclude with a summary of the findings.

The summary, conclusions and implications will be the basis of Chapter V. It will include the following: a summary of the design, the findings, conclusions, implications for education, and implications for future research.

CHAPTER II

REVIEW OF THE LITERATURE

Definition of Writing

The study of language production has previously concerned itself almost exclusively with the oral and visual-manual modes (Kowal & O'Connell, in press), and writing research is considered by many to still be in a preparadigmatic stage (Bereiter & Scardamalia, 1983). Each researcher has his or her own idiosyncratic definition of the term "writing." Hörmann (1976) used the term to refer to the human system of intercommunication conveyed by conventional marks. Murray (1978) defined writing as the process of using language to discover the meaning of experience and to communicate this meaning. On the other hand, Heath (1981) operationally defined writing as the basic skill learned in school; specifically she referred to the capacity to sign one's name and the ability to produce written words in response to a request to do so. For the purposes of this study, writing will refer to both the processes involved in putting words on paper as well as to the completed text.

Currently writing research focuses on its processes, development, and function. Writing research in the field of education naturally revolves around the school essay. The essay is expected to be free of spelling and punctuation errors, and consist of logical, organized

sentences and paragraphs. Researchers caution, however, that although the school essay may be critical for academia, it is not necessarily relevant to later work activities. This is particularly important when one considers students whose own language does not conform to the regulations of standard written English (Hendrix, 1981).

Several researchers have commented that the emphasis on spelling, punctuation, and grammar results in a misdiagnosis in assessment of an individual's educational and intellectual achievement. Yet most researchers agree that the state of one's writing does indeed constitute a valid source of measurement of intellectual achievement (Corbett, 1981). Weiner (1980) has pointed out that:

The most accessible and useful material for diagnostic teaching is a student's own writing. The written production is especially valuable because in the process of transferring thoughts to words, the student learns everything necessary to become a proficient reader and a good writer. The student with deficiencies in both areas welcomes diagnostic teaching. It gives him concrete evidence that improvement is possible if his mistakes are analyzed one at a time, then corrected with the teacher's help until he is able to function independently in monitoring his own work (p. 43).

The essay allows the student to do more than simply list facts; it serves as an example of scholarly performance and in this respect may be a direct measure of educational and intellectual achievement (Coffman, 1971; Klein, Hart, & Frederick, 1968). But the type of writing task requested of the student may influence the quality of his or her writing, and it is in this respect that researchers have concluded that grades administered on written essays indicate little about student learning per se (Cohen, 1973; Gregg, 1983).

Bamberg (1978) compared the high school preparation of 178 University of California college freshmen in regular and remedial

English classes and concluded:

Present concern about the teaching of composition is helpful insofar as it reaffirms the importance of learning to write and the need to teach students to write. However, the concern may be harmful if overly simplistic solutions are adopted in an attempt to improve writing. Successful writers in this study appeared to write better because their high school English programs provided more opportunity to learn. Moreover, the curriculum emphases which contributed most of their superiority were practice in expository writing and instruction in content development and organization (p. 58).

What characterizes "poor" writers in the college level educational system today is not that they do not understand the rules and mechanics of good grammar. Poor writers exhibit underdeveloped composing processes. The processes of composing are rarely explained to students, yet this one factor significantly influences the writer's final product (Pianko, 1979).

Several researchers have agreed that before writing as a skill can be improved, we need to understand the development and psychological processes behind the ability (Kowal & O'Connell, in press; Rosenberg, 1982; Slobin, 1979; Stallard, 1972; Walvoord, 1982). Students who are coached with their writing processes rather than having their essays simply judged, show improvement in the quality of their work (Walvoord, 1982). However, if educators are to instruct students about writing processes, they need an improved understanding of these processes (Kowal & O'Connell, in press). This is further complicated by the issue or topic being discussed in the written composition; unclear instructions, vague issues, and/or lack of an issue all aid in determining the quality of the final product (Della-Piana, Odell, Cooper, & Endo, 1976).

One approach to understanding writing processes has been a physiological one. Osgood and Sebeok (1965) perceived the human nervous system in terms of several levels of organization; they identified these levels as motivational, semantic, sequential, and integrational. These levels of processing, in turn, influence the transformation of thought into action, the action which can then produce a composition. This process of composing is a complex one, involving memory, cognition, language, and psychomotor skills (Glatthorn, 1981). Hays and Flower (1983) have identified four global features of writing which assist in better understanding its cognitive processes: writing consists of distinct processes, these processes are highly embedded, writing itself is a goal-directed process, and writing stimulates the discovery of new goals. Finally, before elaborating on the various theories of writing it is important to remember that the process of composing is a recursive and interacting one, its stages overlap, interact, and influence one another (Glatthorn, 1981).

Composition Processes

The past decade has seen an increased elaboration and understanding of the stages of writing. Generally, writing can be explained in terms of planning, composing and revising. For instance, Pianko (1979) broke down the composing process into seven stages. The first identifiable stage is prewriting; basically this is all of the activity the writer engages in from the time the assignment is received until the first words are put on paper. The second stage is planning. This involves establishing general or

specific (written and/or mental) parameters for the essay. The third stage is composing, and it has three substages, writing, pausing, and rescanning. Composing is what occurs between the time the first words are put on paper and when the writer stops writing. Pianko defines writing as the actual motoric process of putting words on paper; pausing includes those breaks in writing for thinking (filled pauses) or for diversion (unfilled pauses); rescanning is rereading the words, sentences, or paragraphs. Rescanning is not the rereading of the entire essay. Rescanning is that point where revisions are typically made, usually involving a few words or punctuation changes, and the writer might also take some time to contemplate what he/she has written. The fourth stage, rereading, is when the writer rereads the whole essay so as to assess what has been done, to revise and proofread, and if necessary to formulate a conclusion. The fifth stage is stopping: the writer thinks he/she has written everything necessary and desired about the topic for that particular time. Pianko identified the sixth stage as contemplation of the finished product, often for just a short time. Finally, the last stage involves handing in the final product. Pianko has pointed out that for some this is a quick and easy job, while for others it seems to be quite an emotionally laden activity.

Hays and Flower (1980; cited in Burtis, Bereiter, Scardamalia, & Tetroe, 1983) have developed a theoretical model to explain the writing process. They divided writing into three interacting processes: planning, translating, and reviewing. Planning permits the writer to set goals and create a plan for composing a text that

will meet the goals. Planning is seen as consisting of three sub-processes; the writer retrieves relevant information from long-term memory, organizes it, and then uses it to set goals for the text. Hays and Flower have reported that expert writers can be distinguished from inexpert writers not only by the texts they produce but by the kind and amount of planning they use. The more systemized the planning, the better the text. The process of translating utilizes the plan to transform information in the writer's memory and work it into the text. The process of reviewing entails reading and editing the text which is produced via the translating process.

Mosenthal (1983) defined the writing process as the ". . . possible ways a writer may recruit and integrate semantic information from the different meaning sources to produce written text" (p. 33). Here, Mosenthal used the concept of meaning sources to refer to a set of semantic information that could be used to produce a text at any given point in time. According to Mosenthal, this process consists of reproduction, reconstruction, and embellishment. Reproduction is the manner in which the writer produces a text; it involves pulling relevant information from meaning sources. Reconstruction is the process of producing a written text by inferring information from meaning sources. Finally, embellishment refers to the writer's production of inferences which have no identifiable antecedents in previously used meaning sources. A new meaning source has been created.

Further clarification comes from Cooper and Matsuhashi (1983) who discussed the planning of written discourse. They referred to

global discourse plans, the decisions made throughout text production. Initially the writer must identify the purpose of the text and the audience reading the text. The writer must also decide upon a structure. Structure determines the discourse pattern, the semantic role of the sentence, and the abstraction level of the sentence. This will influence and be influenced by the type of discourse used and may be expressive, poetic or transactional in nature. Cooper and Matsuhashi envisioned the writer making decisions regarding sentence plans and production. These are propositional decisions--the wording and presenting of sentences. In conjunction with this, the writer is storing these decisions in short-term memory and executing the graphomotor plan emanating from these decisions.

The cognitive processes which result in writing performance have been described by Hays and Flower (1983); the three main factors are the task environment, the writer's long-term memory and the monitoring processes. They have defined the task environment as the rhetorical problem (topic, audience and exigency) and the text produced thus far by the writer. The writer's long-term memory holds knowledge of the topic, recognition of an audience and his/her writing plans. Finally, the monitoring processes start with planning (organizing, generating and goal setting), then translating, and reviewing (evaluating and editing).

Obviously, these processes are complex and difficult to observe directly. Somehow researchers need to develop methods of generating data from which valid inferences of the processes can be made (Odell, Cooper & Courts, 1978). The cognitive-developmental approach seeks

to explain how the skills of composing develop and how different individuals are able to accomplish various cognitive tasks.

Martlew (1983) has developed such a model of the cognitive processes involved in writing. Initially, the writer becomes aware of the goals of the text to be written, included in the formulation of these goals are the topic, reader, discourse mode and style. Directly influencing the goals of the writer is memory that produces information, and develops strategies applicable to the particular writing task at hand. Memory is indirectly prompted by the external aids available to the writer, such as instructions, procedures, and references. Both memory and goals determine the global and/or local plans decided upon by the writer. These plans then form a reciprocal relationship with the type of linguistic expression decided upon by the writer which include both semantic and syntactic expression. There is also a reciprocal relationship between linguistic expression and text production. In order to produce a text, the writer must incorporate spelling, writing, punctuation and other such skills. Directly influencing memory are plans, linguistic expression, and text production, all of which are cognitive procedures adopted by the writer. The writer must recognize, select, compare, and organize the material for the text. In addition, the writer must also evaluate, edit and revise his/her own writing. Influencing the relationship between goals and plans is cognitive awareness. Cognitive awareness is made up of four elements, the writer must recognize the need to act, realize how to act, sustain the action over time, and integrate it with other procedures.

Now it can be seen that writing does not occur in a vacuum, the text is produced only as the writer searches for and discovers what he/she would like to say. Writing thus becomes the process by which language is used to find meaning from previous experience and then to somehow communicate this meaning (Della-Piana et al., 1976; Murray, 1978). All of the theories of writing start by explaining some sort of a prewriting process; the writer assembles old information and/or gathers new information and proceeds to sort through values and feelings associated with the topic. Writing and rewriting make up the generally accepted actual text composition process.

While the writer is moving through the text composition process, he/she also faces demands from two sources, content and structure. Content demands are those associated with the production of sound ideas, the structure demands are those associated with the development of cohesive sentences. One source of ineffective writing may be a difficulty in satisfying both content and structure demands at the same time. Perhaps the writer becomes overwhelmed by the demands and is unable to proceed with the dissemination of ideas (Glynn, Britton, Muth, & Dogan, 1982).

A logical question that one may ask is: What happens in the composing process of some writers that makes them better than others? Several researchers (Glatthorn, 1982; Stallard, 1972; Walvoord, 1982) have tried to compare the composition processes of skilled writers with those of poor writers and to elaborate on differences between the two groups.

Glatthorn (1981), after an extensive review of the literature,

broke down the composing processes of skilled writers into five categories: exploring, planning, drafting, revising, and sharing. He reasoned that skilled writers believe exploring activities to be useful and helpful and are willing to spend time involved in consideration and contemplation of the topic. Skilled writers are more prone to use note-taking, sketching and diagramming in their planning processes. Skilled writers are less likely to draft their text in language that is like speech; they demonstrate a sensitivity to the reader in their writing, and they spend more time drafting than do unskilled writers. They discontinue writing frequently to rescan, reread, and reflect on the text. Skilled writers, in the process of drafting, are capable of responding to the audience, medium and voice aspects of the rhetorical problem. It is interesting to note that Glatthorn found that skilled writers either revise very little or revise extensively at the sentence and paragraph levels. They are concerned with the content and reader appeal of the text rather than with form. He found that skilled writers viewed revising as a recursive and ongoing process. Finally, Glatthorn found that skilled writers are eager to share their texts, thus providing an opportunity to receive constructive criticism. And finally, the skilled writers see publishing and the dissemination of writing as important.

Glatthorn's review of the composing processes of unskilled writers also considers writing in terms of exploring, planning, drafting, revising, and sharing. Glatthorn found that unskilled writers do not consider exploring important and useful, and as a consequence they spend little time in the activity. Unskilled writers

typically do not make plans before they write, prefer not to use outlines or choose to make an outline after the text is drafted. It appears that they develop limited plans as they write. The drafting process of unskilled writers reflects a written product that imitates spontaneous speech, a lack of concern for the audience, a preoccupation with spelling and punctuation, very little pausing to reflect and reason, and a focus only on the topic, with little concern for "the whole rhetorical problem" (p. 5). So much time is spent correcting spelling and punctuation at the drafting stage that the writer loses sight of topical concerns. There is very little revising, although it does occur at the surface and word levels. Revision stops when the writer feels he/she has not violated any rules. Very often, revising is seen as simply making a neat ink copy. Unskilled writers are not likely to share their writing and do so simply as a means of receiving reassurance.

Based on a sample of 30 seniors in a Virginia public high school, Stallard (1972) identified several behaviors peculiar to good student writers. Fifteen student writers that ranked highest on the STEP Essay Written Test, were termed good student writers and their writing was compared to that of fifteen randomly selected writers from the same class. All of the students were requested to defend their position on some issue in the news. Using videotapes and verbal reports, Stallard reasoned that good student writers spent more time thinking about the text, prewriting, and writing than did other writers. Stallard found that they were also slower writers, writing almost half as many words per minute as their randomly chosen

counterparts. The good writers changed more words as they wrote and these changes were made during intervals of reading and contemplating the text interspersed during the actual writing process. Stallard also found that the good student writers expressed a concern for having a clear purpose in their writing.

Stallard was also able to identify in good student writers four characteristics also found in the texts of a group of randomly selected student writers. He found the students were concerned about revising in order to correct spelling errors, expressed concern for the general mechanics of writing, and demonstrated a lack of concern for a previously determined structure of paragraphs or the content of their texts.

Pianko (1979), employing a remedial versus traditional student classification (based on social class status), found that nonremedial college writers possess a more fully developed understanding of the elements necessary for a well-developed composition than do remedial college writers. Basing her logic on five writing episodes in which the students were to write 400 word essays, Pianko reasoned that traditional college writers recognize style, purpose and getting ideas across; and so in order to better develop their ideas, they spend more time than do remedial college writers in the processes of prewriting, pausing, and rescanning so as to insure clarity and organization. Pianko concluded that better writers (i.e., traditional writers) were better able to "reflect on what is being written" (p. 20), whereas the remedial writers were not able to slow down the process so as to reflect and strengthen their texts.

Developing a reader-based vs. a writer-based prose style is another orientation used to differentiate good writing from bad writing (Flower, 1979). Flower stated ". . .effective writers do not simply express thought but transform it in certain complex but describable ways for the needs of a reader" (p. 19). Reader-based prose is based upon a purposeful attempt to share language and context so as to communicate something from a writer to a reader. Rather than presenting the writer's discovery process, reader-based prose attempts to give the reader an issue-oriented rhetorical text. This is accomplished in the text via its language and structure; it reflects the purpose of the writer's thought rather than its process. In this context, Flower defined good writing as ". . .the cognitively demanding transformation of the natural but private expressions of writer-based thought into a structure and style adapted to a reader" (p. 20).

Writer-based prose, on the other hand, results in a composition by the writer, for the writer. It reflects the writer's own narrative purpose and style without regard for the reader: connections are not clear, privately loaded terms are used, and the reader must follow the writer's own interior monologue rather than a text free of the writer's own idiosyncratic verbiage (Flower, 1979). Additional support for this hypothesis comes from Martlew (1983) who explained that the texts of poor writers read as though the writer were speaking to, rather than writing for, a particular audience. The writer is unaware of the "decontextualized nature of written language" (p. 296). There is little revision nor are referents specified. Martlew has proposed that poor writers may lack an adequate schema and without

automatic writing skills they are incapable of concentrating on the composition itself.

Gregg (1983) found the problem of lack of connection or cohesive ties to be a critical one for college level writers of average mental ability who are underachieving in writing abilities. Martlew (1983) proposed that the writer's constant need to attend to the reader may be the most problematic aspect of writing; it is necessary for the writer to constantly be aware and provide himself/herself with prompts to enforce this vigilance.

Hence, competent writers slowly develop an ability to identify and correct problems in composition. Poor writers, however, appear to retain the writing characteristics of the beginning writer. In addition to a lack of regard for the reader, there may be a lack of progression and integration of writing skills. One can easily identify many possible reasons for this. Lack of motivation or a learning disability, ineffective teaching environments, lack of awareness that a problem exists, habituation in the composing process, are only some of the factors which could possibly correlate with poor writing. Writers who are aware that they are somehow inadequate may lack the skills or resources to improve. Others may find the writing task so complex that they cannot keep all of the rules, goals, and guidelines of the writing task available (Walvoord, 1982).

If the composing process does, indeed, overtax the cognitive abilities of some writers, perhaps it is necessary to develop a means to reduce this strain. Other writers may be aware of the cognitive procedures necessary for writing, but do not recognize the need for

them. Simply having knowledge about writing is not enough; writers need to know when to use this knowledge. By designing instructions incorporating known procedures based upon factors of the writing process, it may be possible to reduce the load carried by the working memory at the time of composition. As proposed by Black (1982), writers need to learn to modify their methods of oral discourse so as to make it appropriate for written production. He suggested that ". . .during conversations people are prompted for the next part of the discourse by the person they are talking to, whereas in written composition people have to generate these prompts themselves. Thus, from this perspective, we would expect that prompting writers with potential discourse elements would facilitate writing. . ." (p. 208). Types of prompts might include directives, sentence openers, and cueing. Black also has pointed out that these types of exercises also provide a means to experimentally validate procedures designed to manipulate the composing process. He then explained that: "Expanding the available repertory of writing facilitation procedures and determining precisely when they are facilitatory [sic] are important tasks for future research" (p. 211).

Somehow, researchers need to be able to assess how deliberate manipulation affects the mental and physical activities that result in the formation of a written text (Bracewell, 1983). Additionally, it is necessary to identify what sequences of behavior produce the written text and how variations of these sequences serve to produce variations in written texts (Perl, 1979). Failure to partake in this activity will result in teachers presenting another method of writing instruction to students who already possess internalized writing

processes. The teacher needs to assist student writers in straightening out and improving defective writing processes. Pianko (1979), in discussing methods of helping students to write better, stated that education must consider changing its focus from evaluation and correcting completed texts to assisting students in expanding and elaborating the stages of their composing processes. Student writers need to be more aware of the dimensions of composition which enable them to "explore ideas, concepts, happenings, emotions, speculations and through which they can evolve a sense of commitment to and understanding about writing" (p. 21). In other words, Pianko proposed that teachers aid students in becoming more reflective writers.

Revision

Revising and rewriting are activities usually considered to be essential to good writing. Revision strategies vary among individuals and there is no particular strategy that emerges as superior. However, the ability to reflect upon what has been written is what is considered to separate good from bad writers (Maimon, Belcher, Hearn, Nodine, & O'Connor, 1981; Pianko, 1979). Sommers (1980) surveyed experienced adult writers in an attempt to compare their revision strategies with those of student writers. In the course of the research Sommers redefined revision as "a sequence of changes in a composition--changes which are initiated by cues and occur continually throughout the writing of a work" (p. 380). The experienced adult writers (20 journalists, editors, and academicians from Boston and Oklahoma City) stated their primary objective in revision is to find form or shape in the written text. They imagine a reader, reading

their text, whose expectations then influence what will go into the final draft. The experienced writers sought to discover or create meaning in the process of their writing. The heaviest concentration or changes was at the sentence level, primarily by addition or deletion. Finally, the experienced writers adopted a holistic perspective with regards to the text and recognized that revision is considered to be a worthwhile activity but one that is frequently overlooked in research and application.

It was not surprising, therefore, that Sommers (1980) found student writer revision strategies to be qualitatively different from those of experienced writers. The 20 students surveyed, freshmen from Boston University and the University of Oklahoma, saw revision as a rewording activity and they believed that most of their problems could be solved through rewording. The students made lexical, but not semantic, changes because economy was their main goal--they worried most about being repetitious. It appeared that the students did what they had been taught to do in a consistently narrow and predictable way. The student writers were unable to see revision as a process; they could not look at their texts from a perspective other than their own. Sommers reported that revision strategies were instructor-based, directed toward a teacher-reader who demands compliance with rules, even if those rules do not apply to the specific problems of the text. The students appeared to lack procedures to aid them in reordering lines of reasoning or to ask questions about the purposes and readers of their texts.

Revision should serve as an opportunity for students to

utilize suggestions for improvement (Bamberg, 1978). By splitting revision into two separate processes, internal and external revision, Murray (1978) hoped to aid educators in better assisting student writers. Internal revision begins with the rereading of the first draft. This reading is to aid students in discovering where the essay's content, form, language, and voice have gone. The writer assesses and outlines what he/she hopes to say. At this point there is only one person in the audience, the writer. Internal revision involves the meshing of content, form, structure, language and voice. Murray identified external revision as the writer's attempt to communicate what was created during internal revision to an audience other than him- or herself. It involves editing and proofreading. By identifying the audience, the writer attends to form, language, mechanics, and style, so as to express thoughts appropriately. Murray hoped that by introducing student writers to revision, the prospect of a second draft would become one of improvement and discovery, and eventually a natural part of the writing process, rather than simply meaning failure on the part of the first draft.

Sommers (1980) has highlighted that revision is a sequence of changes in a written text; these changes occur continually throughout the writing of the text and are initiated by cues. Perhaps student writers fail to recognize the cues present in their text which indicate the need for change. Olson and Torrence (1983) have suggested that student writers lack the ability to adequately connect the content and intent of their essays. Without this connection, students cannot revise nor can they readily evaluate

what has been written.

It appears, then, that student writers may perceive revision as improvement when considered as part of writing, but as flaws when viewed as part of the final text (Kowal & O'Connell, in press). Student writers do not perceive revision as a process of discovery (Della-Piana et al., 1976). There appears to be a failure to incorporate new ideas with those already on paper, an inability to suggest to others the content of a text via structural changes.

Improving Student Writing

The recent surge of interest in the back-to-basics movement has brought about an increased concern for student writing. Testing student writing begins early and often does not end until professional certification. At the college level educators are typically concentrating on assisting students with mechanical and syntactic problems (Freedman, 1979) via drills and homework (Hendrix, 1981). Thus, the writing instructor has become an editor of the final product rather than an assistant of the writing process, content and organization (Walvoord, 1982).

The past decade has seen a tremendous amount of research into the variables which serve to assess and improve writing. Bamberg (1978), in a review of the literature, has pointed out that there are some variables which consistently are able to help students improve their writing skills while others seem to be unable, by themselves, to do so. For instance, the application of transformational grammar, rather than the study of it, results in an increase in fluency. Research in composition has found that instruction and/or

revision in combination with writing practice and applied functional grammar instruction consistently improve writing skill. Arnold (cited in Jerabek & Dieterich, 1975) and Bamberg (1978) found that increasing the frequency of writing or increasing the severity of evaluation does not improve student writing. In fact, the combination of the two was no more effective than infrequent writing and moderate levels of evaluation.

An unpublished study by Cooper, Cherry, Gerber, Fleisher, Copley, and Sarlisky in 1979 (cited in Black, 1982) of the Fall 1979 entering class at State University of New York at Buffalo, attempted to specifically identify the students' writing problems. The authors reported that at the word and sentence levels, the students produced nearly perfect writing. They had shown mastery over standard usage and punctuation. But the students demonstrated great difficulty in producing written texts utilizing adequate connections and relationships from sentence to sentence. The students were also unable to produce examples, anecdotes and details to support their generalizations. The conclusion to be drawn from this study is that students do not necessarily need more drill in the mechanics of writing, but that they need more instruction in the written composition process, specifically at the intersentence level.

Glatthorn (1981), in a review of the literature directed towards administrators rather than educators, found several relationships between certain instructional activities and writing improvement. The study of formal grammar is not related to improvement in writing and may only serve to take time away from writing instruction.

Sentence combining activities do assist students in writing more syntactically mature sentences. Increasing writing frequency without other activities does not improve writing, nor does the type or intensity of teacher evaluation. There is a positive relationship between increases in reading and writing improvement. Prewriting activities, such as exploring and planning, help students learn to write better. Improvements in writing can be brought about by incorporating peer feedback and peer editing. There are no significant differences between positive and negative teacher criticism of student writing. The evidence is in support of revision as a strategy to improve writing (de Beaugrande, 1984; Maimon et al., 1981), although there is no evidence supporting one revision strategy over another.

Citing the conclusion of an M.A. thesis by Underwood completed in 1968, Jerabeck and Dieterich (1975) reported that grades and marginal comments are effective in improving mechanics of student essays, but the content improves more when revisions are required and no grades or marginal comments are used.

Walvoord (1982), in a guide for teachers of all disciplines, has suggested that the manner in which the assignment is presented to the students can also facilitate their writing. Identifying the audience, defining what a successful text looks like, suggesting procedures, and defining and focusing the topic should leave little room for confusion and misinterpretation. Black (1982) has also provided some instructional methods for teachers. For instance, have the students make a list of content words representing ideas that they might want to use in their texts. Provide a list of sentence

opener prompts or perhaps supply a list of prompts that indicate the kinds of discourse elements that are available for them to use.

Flower and Hays (cited in Burtis et al., 1983) have recommended planning as a way of reducing the cognitive strain of writing without sacrificing attention to the requirements of the task. Similar to the notion of planning is the suggestion of Maimon et al. (1981), in an undergraduate writing text, that writing is akin to solving a problem which has more than one solution. They suggested that students begin with a definition and interpretation of the problem facing them. They propose defining the problem in terms of its aim or purpose and its audience. By establishing guidelines such as these for themselves, students reduce the cognitive load they must continually carry while writing the text.

Writing apprehension has also been found to be a factor associated with writing proficiency. Highly apprehensive writers find writing activities unrewarding and perhaps even punishing, they therefore avoid situations that call for writing. Low apprehensive writers, on the other hand, are confident in their writing abilities, tend not to avoid situations requiring writing and frequently enjoy writing (Faigley, Daly & Witte, 1981). Daly and Miller (1975) proposed that when people are placed under cognitive stress (i.e., apprehension), they generally produce less intense communication than when they are not under stress. Therefore, stylistic differences may exist between students who are generally apprehensive about writing and those who are not generally apprehensive about writing. Analyzing the essays of high and low apprehensives, Faigley et al. (1981) found

that the highly anxious writers produced essays that were shorter and less syntactically "mature" or "fluent" than the low apprehensives. They also reported that the written texts of high apprehensives had fewer words, less qualification and lower intensity as compared to the texts of low apprehensives. In addition, the texts of high apprehensives were evaluated less positively than those written by low apprehensives. The researchers pointed out, however, that writing apprehension cannot be assumed to be a causal factor of poor writing nor can poor writing be assumed to be a causal factor of high apprehension; apprehension and performance most likely reinforce one another, producing a bidirectional relationship. The authors suggested that, given their differential performance in writing, different instructional materials and methods may help highly apprehensive writers.

Perl (1979) studied the composing process of five unskilled community college writers and found that they all displayed consistent and internalized, albeit ineffective, writing process. Strategies such as creating an association to a key word, focusing in or narrowing down the topic, dichotomizing and classifying were all effective in helping the students to improve their writing. Perl found that these types of activities can be taught in a relatively brief period of time. She also reported that the development and clarification of ideas are facilitated once students put their ideas onto paper; this seems to provide an opportunity for the students to reflect upon, change and develop their ideas further. However, it is unclear whether these methods, as well as the others discussed

in this section, help to maintain improvement in student writing over time.

Methods of Writing Evaluation

Indirect measures of writing abilities are typically standardized tests designed to assess students' writing proficiency; they allow many students to be tested at one time, are objective, easily scored and evaluated. However, they are incapable of assessing how well a student actually composes in and outside the classroom. There are doubts as to the ability of the tests to truly measure students' abilities to produce written texts. Daly (1978) has pointed out that ideas generated by the writer ". . . are not and indeed probably could not be assessed through an objective testing procedure" (p. 13).

Della-Piana et al. (1976) reported that the results of standardized tests do not appear to be highly related to out-of-class writing success. Standardized tests, specifically the McGraw-Hill Basic Skills System Writing Test, Sequential Test of Educational Progress: Writing, and the Missouri College English Test, are designed in the multiple choice format and request that students choose from previously established alternatives. The process of creating a text requires that the writer produce his or her own alternatives, and then choose the best one. Standardized tests cannot assess the skill of conceiving and planning a text containing several paragraphs. Hendrix (1981) has also stated that assessment does not serve to improve writing. When writing moves from being a means of learning, expression and improvement to testing, it simply results in another method of sorting and certifying students (de Beaugrande, 1984).

By asking students to read a written passage and make the relatively minor alterations necessary to transform the passage into an acceptable finished draft, the tests fail to measure the full range of activities necessary for effective writing (Della-Piana et al., 1976). The standardized tests fail to measure how well the student writer is able to develop choices, such as developing a sense of purpose in the text and developing a sense of relationship to the reader.

The use of direct methods (i.e., the production of essays) to evaluate writing have been strongly criticized by arguments claiming a lack of objectivity and fairness. However, research has shown that ". . . human readers can make reliable judgments and develop reliable descriptions of crucial aspects of actual pieces of writing. . . ." (Della-Piana et al., 1976, p. 36). Reliability of .90 and above can be achieved with specific and clear scoring criteria and a short period of training (Della-Piana et al., 1976; Powills, 1979).

As with any other skill, there are several ways to approach assessment of student writers. Odell (1981) has indicated that there are three primary questions to be asked in evaluation. First, are individual students improving as writers? Second, is the program in use helping large groups of students? And third, which students have not achieved at least minimum competence in their writing?

One direct method of assessment which has proven itself capable of answering these questions is "holistic essay scoring" (Gregg, 1983; Warshauer & Calfee, 1983). This method assigns each essay a numerical score based on the reader's opinion of the overall quality of the

essay. The primary purpose behind such a method is to score the entire essay, not just the frequency and types of errors. An assumption here is that errors in and of themselves do not reflect the writer's competence. Research by Freedman (1979) supports the use of this method. Freedman sought to determine what characteristics of student writing most influence teachers' evaluations. Based on evaluations by twelve teachers on the staff at Stanford University's freshman English program, she concluded that the two most important characteristics were the essays' content and organization, which are amenable to direct assessment, while sentence structure and mechanics were shown to be less important.

When the holistic scoring procedure is used for experimental purposes one needs to insure that the method provides reliability and adequately assesses the experimental intervention. Coffman (1971) as well as Klein and Hart (1968) identified three factors which need to be taken into account when assessing the reliabilities of holistic essay examinations: 1) Do different readers (scorers) tend to assign the same grade or score to the same paper; 2) Does a single reader assign the same score to the same paper on different occasions; 3) Do these reliabilities tend to increase as the essay topic allows for greater freedom of purpose. Intrareader variability can most likely be attributed to the relative standard for scoring different essays, the general scoring standards, and the variability of the ratings themselves. Hence, there is a need for at least two readings by each reader of each essay; the sum of two readings will be more reliable than a single reading score (Coffman, 1971; Della-Piana et al., 1976).

Diederich (1966) has recommended that the reliability of the cumulative total for each essay should reach (or nearly reach) .90, which is often cited as the minimally acceptable reliability coefficient for individual measurement. Interreader reliability can be assessed by tabulating a coefficient of correlation from the scores of different readers across essays. Inter- and intrareader variability can also be examined via the reliability of assigned scores on essays written on different occasions (Klein & Hart, 1968).

Frequency of Micro-elements and Essay Writing

At first glance, it may appear that the use of statistical theory in the analysis of essay writing is inappropriate. But Kowal and O'Connell (in press) have pointed out that this is not only a legitimate activity, it also allows the researcher to confidently make generalizations about the style of the writer based on a sample of his/her work. Statistical procedures permit the researcher to present ". . . compact descriptions of quantifiable features of style" (p. 15). The micro-elements of a written essay which are amenable to statistical analysis include the length of words, sentences and paragraphs. Rhetorical devices, syntactic constructions, and vocabulary can also be counted and their frequencies determined. However, Kowal and O'Connell also have stressed that the distribution of quantifiable features of written texts is highly skewed, therefore not conforming to the normal bell-shaped curve so often assumed in statistical theory.

These quantifiable aspects of the written essay are indicative of several writer-based characteristics. Total length of the essay

is one indication of the invention skills of the writer (Faigley et al., 1981). Use of words which occur frequently in the language will make a text more easily read. Short words, which are typically more frequently used, make a text more readable than long words. Shorter sentences are typically more comprehensible than longer ones (Hörmann, 1979). However, a text which is confined to a very general vocabulary is considered to be lexically restricted. Hörmann (1979) postulated that the lexically restricted text emerges when the writer perceives the writing situation to be overly stressful (refer back to Faigley et al. and their ideas about the "apprehensive" writer); the writer experiences rigidity of thought and a reduction of creative ideas. Lexical restriction also emerges as a function of effort. The writer who attempts to maintain as low an average level of exertion as possible will produce a text with little vocabulary variability. To assess lexical restriction, Hörmann proposed the token-type ratio. The token-type ratio is derived by taking the number of different words in the text over the total number of words in the text. Hörmann concluded that there is a significant relationship between the token-type ratio of a text and its lexical complexity, and that the less frequently the words in a text are used in everyday language, the more complex the text.

Summary

Writing, in the context of this paper has been defined as both the processes involved in putting words on paper as well as the completed text. It appears that once an individual has started his or her post-secondary education, writing processes have become

internalized and remain relatively consistent over time. In addition, students at the college level are typically required to take one or two writing courses their freshman year. The skills they have developed in high school and in the college level writing classes are expected to provide them with the proficiency they need to write throughout their college careers.

Recently both experts in writing and educators have expressed dissatisfaction with the quality of college student writing. Combined with declining scores on a variety of aptitude and achievement tests, this decrease in the quality of student writing has prompted researchers and educators to explore methods of helping students to improve their writing. Most methods developed to help students improve their writing have hoped to intervene at the process level--in the planning, writing and rewriting stages. Some approaches try to teach students new processes; however, these approaches simply impose new methods on preexisting, internalized methods and their effectiveness is questionable. Other approaches deal with improving weak areas in the internalized process. At the planning stage, students may be asked to produce an outline or thesis; at the writing stage students may be asked to participate in grammatical exercises; at the rewriting stage students may be asked to revise their essays. To date there have been few empirical studies assessing the effectiveness of these methods.

Assessing the various interventions has proven to be a difficult task. Objective tests fail to adequately assess text production and the students' final product. Holistic essay scoring, with its

tightly controlled criteria, has proven itself amenable to student essay measurement. However, as of the writing of this study, the value of holistic essay scoring in assessment of improvement in the writing of a college level assignment has not been verified. In addition to holistic essay scoring, the frequency of micro-elements in essays has provided a method of quantifying style. It would be of interest to researchers if these two methods, holistic essay scoring and the frequency of micro-elements, were to provide us with a method of assessing improvement in student writing.

CHAPTER III

METHODOLOGY

Subjects

The original population from which the sample was derived consisted of 47 undergraduate students in a General Psychology course offered at Loyola University of Chicago during the Spring 1984 semester. Four students dropped the course, six students later withdrew, four failed to complete the assignment on time, and two failed to provide personal information, all were subsequently dropped from the list of potential subjects. This left a sample of 31 (65.96%) students who participated in the experiment from start to finish.

The sample consisted of 12 males and 19 females, the average age was 18.84 years. There were 19 freshmen, 7 sophomores, 3 juniors, and 2 seniors. Seven of the students did not speak English as a first language. The experimental and control groups were determined in the following manner: The article to be read and reviewed or critiqued by the students, "In the Matter of Man" by Alan L. MacKay (1984), was stapled to the essay instructions; control and experimental instructions were alternated in one pile with the article to be read. The students were told to take one collated packet (article and instructions). This procedure resulted in a control group, $N = 16$, and an experimental group, $N = 15$.

The control group consisted of seven males and nine females, the average age was 19.06. The experimental group consisted of five males and ten females, the average age was 18.60. It was originally planned to conduct statistical analyses by males and females in each group; however, the cell sizes obviously would have been too small to make confident inferences, so it was decided to combine the two sexes and conduct the analyses only by group.

Materials

The students in a General Psychology class were requested to read "In the Matter of Man" (MacKay, 1984). The article was photocopied for each student and extra copies were made available. Attached to each copy of the article was a set of instructions and a request for some personal information. The control group instructions read: "Read the article 'In the Matter of Man' written by Alan L. MacKay. Your assignment is to hand in a one-page typed review or critique of the article." The experimental group received the instructions along with five clarifying points. The additional instructions included the following:

1. Imagine that you are writing this review or critique for a friend of your own age who has not read the article.
2. Consider using definitions, criticisms and/or evaluations in your review or critique.
3. On the back of this page, write down one key word from each of the paragraphs you have written.
4. On the back of this page, write down the topic sentence from each of the paragraphs you have written.
5. Read your rough draft aloud to a friend and ask

him or her for corrections (this will help you identify mistakes in grammar and meaning).

Four weeks after the first essay was turned in by the students, they received a second set of instructions requesting that they either rewrite or revise their essays (see Appendix A). Both the control and experimental groups received the same instructions.

On the last day of class all of the students were given a brief questionnaire to complete. This questionnaire consisted of the following six items: 1) On a scale of 1 to 7, how would you rate the quality of your writing? 2) Name the one best thing that you could do to improve your writing? 3) Recall the essays you wrote about "In the Matter of Man": which do you think was better? 4) Which essay did you spend more time on? 5) Consider the essay that you think was better; list the things that you like about it; and 6) Consider the essay that you think was not as good; list the things that you did not like about it.

Design and Procedure

Based on a review of the literature, it was decided to incorporate several cues into an assignment so as to assess their effectiveness in facilitating student writing. Approximately half of the students in a General Psychology class received an assignment with cues and the other half received an assignment without cues. The essays written in response to these assignments determined the initial assessment of writing abilities. All of the students were then requested to rewrite or revise these first essays. This second set of essays determined the final assessment of writing abilities.

Use of two methodologies to assess the first and revised essays permits a comparison not only of groups and instructional

intervention, but also of the methodologies themselves. Holistic essay scoring has been shown to be a valid and reliable procedure when essays are written and rated in tightly controlled situations (see Powills, 1979, for a review). Limitations of prewriting activities, writing time and assignment type are several such ways of insuring control. Use of the holistic essay scoring procedure to assess its sensitivity in the realistic setting of the college assignment further tests its reliability. Its sensitivity can also be compared to the sensitivity of micro-elements (number of words used in the essay, syllables per word, and token-type ratio).

Preparation of Essays for Analysis

The students' essays used for this study were collected and photocopied. The originals were returned to the students and the copies were retained for analysis. Anonymity of the student writer was insured by blacking out any names, dates, or other identifying marks. The essays were given a numerical code from a list of random digits.

Preparation for the reading session began with a sorting of the essays. There were four groups of essays: first essays written by the control group, first essays written by the experimental group, second (revised) essays written by the control group and second (revised) essays written by the experimental group. There were two copies of each of these. All of the essays were then randomly sorted into eleven manilla folders, with the qualifications that duplicate essays be in separate folders and first essay-revised essay counterparts be in separate folders. Essays of students not used in

the analysis (i.e., students who withdrew from the class or incorrectly responded to the assignment) were retained; they were used in the practice session prior to the reading session.

Another copy of each of the essays (both the first and revised) was used to collect micro-elements as explained above. First, the number of words of each essay was counted and recorded. This will be used to assess the inventiveness of the writer. Then the number of syllables of each was counted and recorded. The syllable per word count was determined by dividing the number of syllables by the number of words in each essay. This will aid in assessing the complexity level of the words used by the students. The token-type ratio was a ratio of the number of different words in each essay to the total number of words. This will determine the lexical complexity of the essay. These raw data appear in Appendix B.

The Reading and Scoring Session

Two graduate students from Loyola University of Chicago were used to read and score the essays. Initially, the rationale behind the use of holistic essay scoring was explained. Readers were told that they were to score each essay on the basis of the overall impression that it makes on them. Special emphasis was not to be placed on any one part of the essay; semantics, syntax, spelling, content, style, grammar, and punctuation were to be taken as a whole. Six essays not used in the analysis were used for practice.

The readers were told that students in General Psychology were instructed to read an article from Science magazine and review or critique it (the readers did not read the article). They were then

told to quickly read the first two essays and assign a score from one to four (one being the lowest score possible, four being the highest). It was explained to the readers that they first needed to decide if the essay ranked in the upper or lower half of the group. Then they needed to decide if the essay ranked highest or lowest in that subgroup. The readers were then asked to share their scores and discuss why they gave each essay their respective scores. They were then instructed to read the next four essays and compare scores. This practice session was designed to establish a group rating consensus for the reading-scoring session. It was stressed that the readers should continue to conform to the standards established by the group during the reading and scoring.

Approximately 140 score sheets were given to each reader (see Appendix A). Use of the score sheets was explained: each reader was to put his or her name or initials on the line marked "Reader Name"; the essay number, which appeared in the upper left-hand corner of each essay was to be printed on the line marked "Essay Number"; after the essay was read they were to put their score on the line marked "Score". They were then given one of the eleven manilla folders containing the essays to be read and scored. They were then instructed to sign their names on each folder as they read the essays, so that they would not read the essays in that particular folder more than once.

The readers were told to take rest breaks between folders, to get up and walk around. Snacks and refreshments were provided to help combat fatigue. After the twelve essays in each folder were read, the readers were told to put the completed folder in the middle of the

table and take another. Periodically, completed score sheets were collected and scores recorded (see Appendix A).

Dependent Measures

The data collected for this study included first and revised essay scores assigned by independent readers, number of words, syllables per word, and token-type ratio. Each paper was read and scored by two different readers (graduate students in the Department of Psychology at Loyola University of Chicago) twice during a reading session. The two scores given to each essay by each reader were correlated to determine intrarater reliability. The two scores assigned by each reader were summed to derive a total essay score. Intrareader reliability coefficients were determined by correlating these two scores. Interreader reliability coefficients were determined by correlating total scores between readers. A grand total for each essay was obtained by summing the two total scores from each reader.

Concurrent validity was determined by correlating total essay scores with grades in Writing I and General Psychology. Twenty-four of the thirty-one subjects provided their grades in Writing I and all of the subjects' grades in General Psychology were made available by the instructor.

Questions To Be Answered

The purpose of this thesis is to determine whether intervention in the classroom, outside of the English department, at the college level can aid students in improving their writing. Although it is beyond the scope of this research to hypothesize as to what intervention

works best, it is hoped that by critically assessing one procedure, further hypotheses can be developed and theory tested so as to assist students of composition. There are two primary questions to be answered here: Does providing students with explicit guidelines (vs. providing one general statement) help them to write better essays? And, does the instruction "rewrite or revise your essay" prompt students to critically assess and therefore improve their essays? Insight into the answers of these two questions will be provided by various levels of statistical analyses.

Analysis will begin with a statistical validation of the holistic scoring method. One assumption underlying holistic essay scoring is that for any group of randomly chosen individuals, it is expected that the distribution of assigned scores will resemble that of the normal bell-shaped curve (Diederich, 1966). Did the distribution of assigned scores of the students in this sample resemble a normal curve? Stated in the form of a testable hypothesis:

H1: The distribution of scores for each of the four readings and summed across first essay, revised essay, control, and experimental groups, should approximate a normal distribution.

As was stated earlier, holistic essay scoring has been shown to be a reliable method of assessing student writing. Therefore, it is expected that different readers (scorers) would assign the same score to the same essay and it is expected that a single reader would assign the same score to the same essay on different occasions (Coffman, 1971; Diederich, 1966; Klein & Hart, 1968). Four hypotheses will be tested

to verify these two assumptions:

- H2a: The intrareader reliabilities reflect consistency in scoring;
- H2b: The interreader reliabilities reflect consistency in scoring;
- H2c: The number of essays with discrepant assigned scores by each of the readers is less than the number of essays with nondiscrepant scores;
- H2d: The number of essays receiving the same total score by each of the two readers is greater than the number of essays not receiving the same total score.

The concurrent validity of a measure is its ability to distinguish between individuals who are known to differ (Kidder, 1981). The concurrent validity of the holistic scoring method will be determined by correlating total essay scores with grades received in Writing I and General Psychology. An assumption here is that the grades assigned in Writing I and General Psychology do indeed differentiate these students. Stated in hypothesis form:

- H3a: There is a positive correlation between grades in General Psychology and grades in Writing I;
- H3b: There is a positive correlation between grades in Writing I and scores of the first essays;
- H3c: There is a positive correlation between grades in Writing I and scores of the revised essays;
- H3d: There is a positive correlation between grades in General Psychology and scores of the first essays;
- H3e: There is a positive correlation between grades in General Psychology and the scores of the revised essays.

Statistical analysis to assess growth in writing ability between

groups and over time will begin with the first set of essays turned in by the students. Were there qualitative differences between the essays written by students given explicit guidelines (experimental group) and students given a general statement (control group) to respond to? Are there significant quantitative differences between these two groups? Specifically, will statistical analysis uncover differences in length, as measured by number of words? Are there significant differences in syllables per word used by students in each of the groups? Is there a token-type ratio difference between the two groups? Stated in terms of empirically testable hypotheses:

H4a: The obtained first essay total scores of the experimental group students will be higher than those of the control group;

H4b: The first essay length, as measured by the number of words, of the experimental group students will be longer than that of the control group;

H4c: The syllables per word ratio of the first essays of the experimental group will be higher than that of the control group;

H4d: The token-type ratio of the first essays of the experimental group will be higher than that of the control group.

The next set of questions to be reviewed in assessing differences between the groups will be explored in the same manner as above. Were there quantitative differences in the revised essays written by the experimental group (students given explicit guidelines) and the control group (students given a general statement to respond to)? Specifically, will statistical analysis find differences in length, as measured by number of words? Are there significant differences in

the syllables per word used by students in each of the groups? Is there a token-type ratio difference between the groups? In other words, did the effect of differential instructions carry over? Here a main effect of group (experimental or control) is expected, but an interaction is not. Stated in terms of empirically testable hypotheses:

H5a: The obtained revised total essay scores of the experimental group students will be higher than those of the control group;

H5b: The revised essay length, as measured by the number of words, of the experimental group will be longer than that of the control group;

H5c: The syllables per word of the revised essays of the experimental group will be higher than that of the control group;

H5d: The token type ratio of the revised essays of the experimental group will be higher than that of the control group.

The first essays written and their revised counterparts will be compared. Did students with explicit instructions improve qualitatively over time? Did students without explicit instructions improve qualitatively over time? Were there qualitative differences in the way that the two groups improved over time? In other words, a main effect of group (experimental or control) is again expected, but an interaction between the two groups is not expected. Stated in terms of an empirically verifiable hypothesis:

H6a: The obtained essay total scores will reflect improvement in writing ability over time within both the control and experimental groups.

Finally, several exploratory questions to assess the construct validity of holistic essay scoring and assessment of the micro-elements will be asked. Assuming that the two measures, holistic essay scoring and measurement of micro-elements, are reliable, a relationship between the two may indicate convergent validity (Kidder, 1981). However, if the two measures prove to be unreliable, their validity will be limited (Brown, 1983). Since holistic essay scoring and measurement of micro-elements have been shown to be reliable indicators of students' writing, convergence of the two methods may demonstrate evidence of construct validity. The scores assigned by the readers and the micro-elements will be compared. Is there a positive correlation between scores assigned by readers and the number of words in the essay? Is there a positive correlation with the syllables per word ratio? Is there a correspondence with the token-type ratio? Based on the students' ratings of their own writing, are their ratings comparable to scores given by the readers? What were the things students thought they should do to improve their writing? How do their answers correspond to the information presented in the research?

H7a: There is a positive correlation between scores on the first essays and their corresponding number of words;

H7b: There is a positive correlation between scores on the first essays and their corresponding syllables per word;

H7c: There is a positive correlation between the scores on the first essays and their corresponding token-type ratio;

- H7d: There is a positive correlation between scores on the revised essays and their corresponding number of words;
- H7e: There is a positive correlation between scores on the revised essays and their corresponding syllables per word;
- H7f: There is a positive correlation between scores on the revised essays and their corresponding token-type ratio.

Data Analyses

The statistical analyses for this study included between means t-tests, univariate analyses of variance, doubly multivariate analyses of variance, and profile analysis. Descriptive statistics included means, standard deviations, and Pearson product moment correlations.

The intrareader reliabilities, interreader reliabilities, and relationships between grades in General Psychology, Writing I, and first and revised essays will be assessed with Pearson product moment correlations. In addition, the scores assigned to the essays and the corresponding micro-elements will also be assessed with Pearson product moment correlations. Comparisons between first and revised essays between and within control and experimental groups will be accomplished via a doubly multivariate analysis of variance. Analysis of variance is used to test the hypothesis that group means of the dependent variable (in this study, assigned scores on the essays, number of words, syllables per word, and the token-type ratio) are significantly different. While multivariate analysis of variance tests the hypothesis that group means of more than one dependent variable are significantly different. A doubly multivariate analysis

of variance assesses two or more dependent variables measured at two or more points in time (Winer, 1971). All of the statistical analyses will be conducted through the SPSSx (1983) statistical package.

Due to the large number of hypotheses being tested a .01 level of probability will be required for significance so as to decrease the chances of spurious statistical significance.

CHAPTER IV

RESULTS AND DISCUSSION

Descriptive Summary Statistics

Each student essay, whether the first or its revised (rewritten) counterpart, was read twice by two readers. Each essay was assigned a score of 1 to 4, with 4 representing the highest score possible. The total score for each essay was the sum of both readers' two scores for the essay. The total maximum possible score was 16, while the total minimum score possible was 4. The distribution of total scores for the control group and the experimental group by essay is shown in Figure 1. The distribution of scores for the control group and the experimental group across essays and readers is shown in Figures 2 and 3. The first experimental hypothesis (the distribution of scores for each of the four readings and summed across first essay, revised essay, control and experimental groups, should approximate a normal distribution) was not supported.

Table 1 shows the mean first essay total scores for the experimental group to be essentially the same as those obtained by the control group. This table also shows the mean revised essay total scores for both groups to be somewhat higher than those obtained on the first essay, with the experimental group again showing approximately the same scores as the control group.

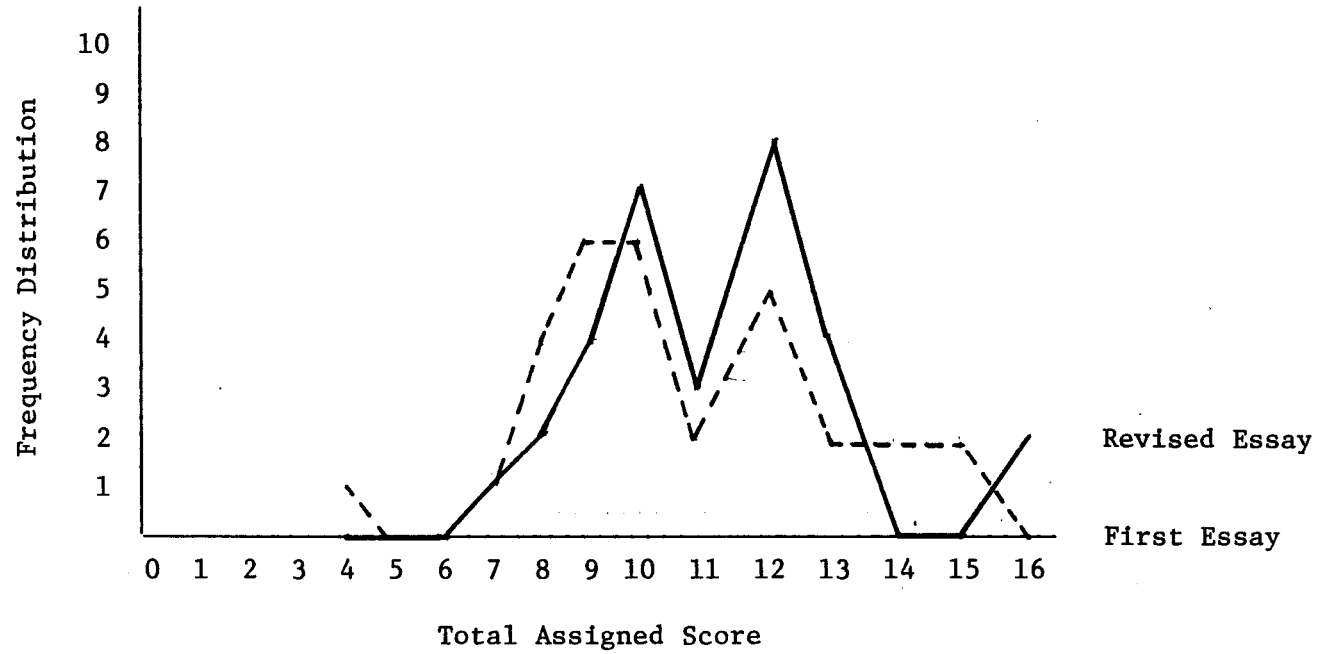


Figure 1. Distribution of Total Assigned Scores by Essay (First and Revised)

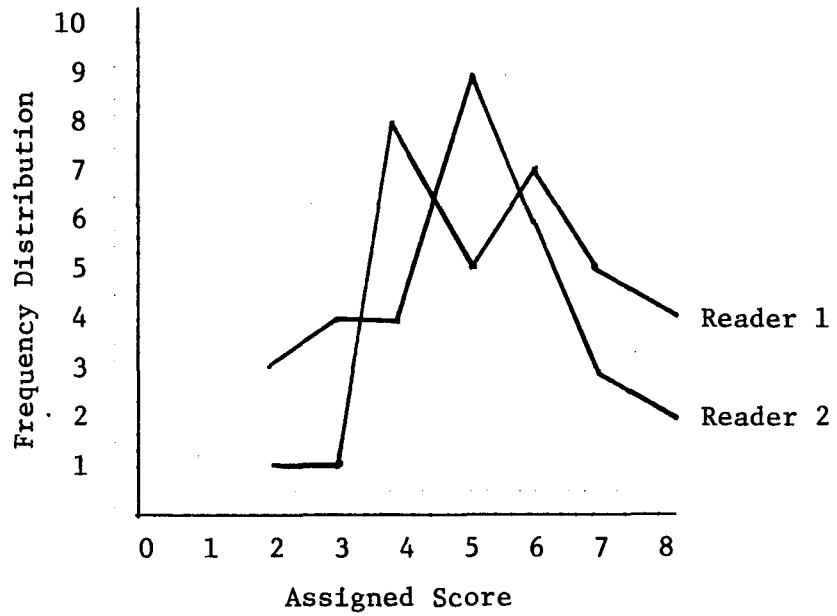


Figure 2. Distribution of Assigned Scores for First Essay by Reader 1 and Reader 2.

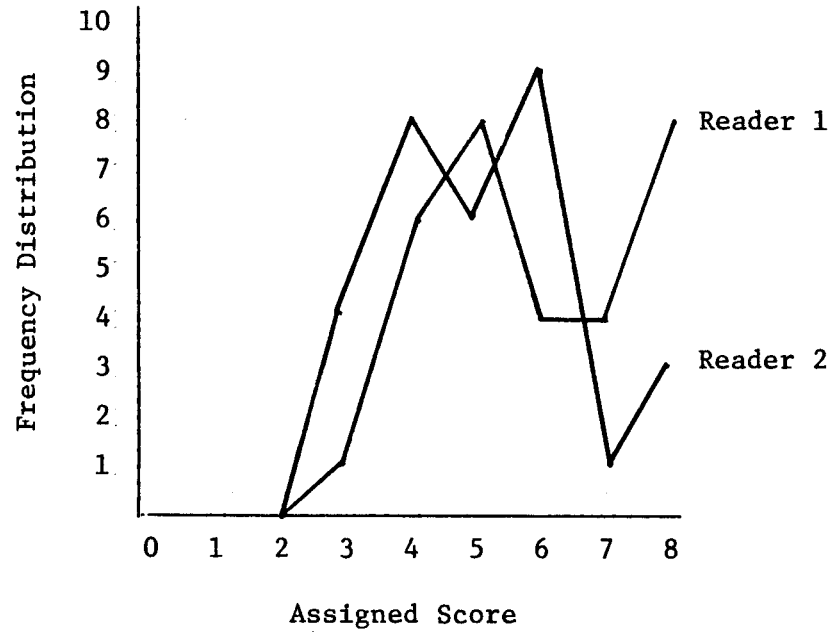


Figure 3. Distribution of Assigned Scores for Revised Essay by Reader 1 and Reader 2.

Table 1

Mean and Standard Deviation of Reader 1 and Reader 2 Assigned Scores and Total Scores for Control (N = 16) and Experimental Groups (N = 15) on the First and Revised Essays

| Group | Reader 1 | | Reader 2 | | Total Scores | |
|----------------------|-----------|-----------|-----------|-----------|--------------|-----------|
| | \bar{X} | <u>SD</u> | \bar{X} | <u>SD</u> | \bar{X} | <u>SD</u> |
| First Essay | | | | | | |
| Control | 5.44 | 1.26 | 5.06 | 1.57 | 10.50 | 2.10 |
| Experimental | 5.60 | 1.92 | 4.73 | 1.79 | 10.33 | 2.90 |
| Revised Essay | | | | | | |
| Control | 5.75 | 1.48 | 5.31 | 1.54 | 11.06 | 2.02 |
| Experimental | 6.07 | 1.75 | 4.93 | 1.39 | 11.00 | 2.24 |

A one-way analysis of variance was conducted to determine if a statistically significant difference existed among the means of the first and revised total essay scores by control and experimental groups. The experimental hypothesis that the groups differed with respect to scores obtained on the first essay was not confirmed, $F(1,29) = .026, p > .05$. The hypothesis that the groups differed with respect to scores obtained on the revised essay also was not supported, $F(1,29) = .042, p > .05$.

Table 2 shows the mean first essay number of words, syllables per words, and token-type ratio for the experimental group to be somewhat higher than those obtained by the control group. This table also illustrates that the mean revised essay number of words used by both groups is lower than that of the first essay, while the mean syllables per word remained the same. The mean token-type ratio of the experimental group diminished while that of the control group showed no change.

A doubly multivariate analysis of variance was conducted to assess micro-element differences between the groups on the first essays written and their revised counterparts. The doubly multivariate analysis of variance results are presented in Table 3. Overall there was no main effect of group membership (control or experimental), $F(1,29) = 1.02, p > .05$. Nor was there a main effect of essay (first or revised), $F(1,29) = 1.09, p > .05$. The hypothesis that the experimental group would use more words than the control group in the first essay was not confirmed. The hypothesis that the experimental group would use more syllables per word in the first essay was not

Table 2

Mean and Standard Deviation for Number of Words, Syllables per Word and Token-Type Ratio for the Control and Experimental Groups Across Essays (First and Revised)

| Group | Number of Words | | Syllables per Word | | Token-Type Ratio | |
|---------------|-----------------|-----------|--------------------|-----------|------------------|-----------|
| | \bar{X} | <u>SD</u> | \bar{X} | <u>SD</u> | \bar{X} | <u>SD</u> |
| First Essay | | | | | | |
| Control | 341.1 | 86.7 | 1.57 | .10 | .363 | .069 |
| Experimental | 351.3 | 87.5 | 1.60 | .09 | .376 | .074 |
| Revised Essay | | | | | | |
| Control | 314.4 | 76.8 | 1.54 | .10 | .336 | .066 |
| Experimental | 299.4 | 60.2 | 1.60 | .11 | .385 | .061 |

Table 3

Doubly Multivariate Analysis of Variance Results Between Control
and Experimental Groups for Micro-Elements (Number of Words,
Syllables per Word, and Token-Type Ratio) on First and Revised Essays

| Essay | Micro-Element | <u>d.f.</u> | <u>F</u> | <u>p</u> |
|---------|--------------------|-------------|----------|----------|
| First | Number of Words | 1,29 | .009 | .924 |
| First | Syllables per Word | 1,29 | 1.998 | .168 |
| First | Token-Type Ratio | 1,29 | 2.078 | .160 |
| Revised | Number of Words | 1,29 | 8.410 | .007 |
| Revised | Syllables per Word | 1,29 | 1.065 | .311 |
| Revised | Token-Type Ratio | 1,29 | .766 | .389 |

supported. The hypothesis that the experimental group would produce a higher token-type ratio than the control group in the first essay was also not confirmed.

The hypothesis that the groups differed with respect to the number of words used in the revised essay was confirmed ($p = .007$): The experimental group used less words on their essays ($\bar{X} = 299$) than did the control group ($\bar{X} = 314$); however, the direction of this finding is opposite to the predicted one. The hypothesis that the groups differed with respect to the syllables per word in the revised essays was not supported. Nor was the hypothesis that the groups differed with respect to the token-type ratio used in the revised essay supported.

Intrareader Reliability

Klein and Hart (1968) have recommended that the reliability of the total scores for each essay reach (or nearly reach) .90, which is often cited as the minimally acceptable reliability coefficient for individual measurement. In accordance with this recommendation, this study also elected to consider .90 as the minimally acceptable inter- and intrareader reliability coefficient.

In this study two readers read both an original essay and its revised (or rewritten) counterpart twice during a reading and scoring session. The rationale used for two readings of each essay was that the sum of two readings would be more reliable than a single reading score (Coffman, 1971; Della-Piana et al., 1976). The intrareader reliabilities are Pearson product moment correlation coefficients.

Table 4 illustrates the intrareader reliability coefficients for

Table 4

Intrareader Reliability Coefficients for Reader 1 and Reader 2
by Control (N = 16) and Experimental (N = 15) Groups Across
First and Revised Essays

| Reader | Group | Essay | Correlation Coefficient |
|----------|--------------|---------|-------------------------|
| Reader 1 | Control | First | .43 |
| Reader 2 | Control | First | .34 |
| Reader 1 | Experimental | First | .74** |
| Reader 2 | Experimental | First | .44 |
| Reader 1 | Control | Revised | .47 |
| Reader 2 | Control | Revised | .35 |
| Reader 1 | Experimental | Revised | .66* |
| Reader 2 | Experimental | Revised | .26 |

* $p < .01$

** $p < .001$

the two readers broken down by group and essay. Testing the experimental hypothesis that the intrareader reliabilities would reflect consistency in scoring, combining control and experimental groups, the correlation coefficients for Reader 1 were $\underline{r} = .63$ and $\underline{r} = .57$ for the first and revised essays, respectively. Combining the control and experimental groups, the correlation coefficients for Reader 2 were $\underline{r} = .38$ and $\underline{r} = .29$, respectively for the first and revised essays, neither of which were statistically significant. None of these correlations approach .90 as suggested by Klein and Hart.

Table 5 illustrates the number of discrepant assigned scores by the two readers. The experimental hypothesis that the number of essays with discrepant assigned scores by each of the readers is less than the number of essays with nondiscrepant scores, clearly was not supported. There were a total of 65 discrepant scores compared to 58 nondiscrepant scores among the two readers. Therefore, use of the holistic essay scoring method in this study failed to achieve the minimum acceptable intrareader reliability.

Interreader Reliability

Interreader reliability was assessed by tabulating a Pearson product moment correlation coefficient from the total scores assigned by the two readers. Table 6 shows the reliability coefficients of the total scores assigned by Reader 1 and Reader 2.

The interreader reliability coefficients of the total scores assigned by both readers did not support the experimental hypothesis that the interreader reliabilities would reflect consistency in scoring. The coefficients ranged from $\underline{r} = -.11$ to $\underline{r} = .22$, obviously

Table 5

Discrepant and Nondiscrepant Assigned Scores by Control and Experimental Groups, Across First and Revised Essays, and by Reader 1 and Reader 2

| Group | Reader | Discrepant Scores | Nondiscrepant Scores |
|---------------|--------|-------------------|----------------------|
| First Essay | | | |
| Control | 1 | 7 | 9 |
| Control | 2 | 11 | 5 |
| Experimental | 1 | 7 | 8 |
| Experimental | 2 | 7 | 8 |
| Revised Essay | | | |
| Control | 1 | 8 | 7 |
| Control | 2 | 10 | 6 |
| Experimental | 1 | 8 | 7 |
| Experimental | 2 | 7 | 8 |
| | Total | 65 | 58 |

Table 6

Interreader Reliability Coefficients of Total Essay Scores
for Reader 1 and Reader 2 by Control (N = 16) and Experimental
(N = 15) Groups Across First and Revised Essays

| Group | Essay | Correlation Coefficient |
|--------------|---------|----------------------------|
| Control | First | .09 |
| Experimental | First | .22 |
| Control | Revised | -.11 |
| Experimental | Revised | .00 |

not even approaching the .90 as suggested by Klein and Hart. Combining the control and experimental groups, the correlation coefficients of $\underline{r} = .16$ (first essay) and $\underline{r} = .23$ (revised essay) also were short of the recommended .90.

The experimental hypothesis that the number of essays receiving the same total score by each of the two readers is greater than the number of essays not receiving the same total score was not confirmed. Reader 1 and Reader 2 matched only one out of sixteen total scores for the first essay of the control group. For the first essay of the experimental group, they matched four out of fifteen. For the revised essay of the control group, the two readers matched five out of sixteen scores, and finally for the revised essays of the experimental group they matched only two of the fifteen total scores. It therefore must be assumed that this study failed to achieve a minimally acceptable level of interreader reliability.

Differences Between Control and Experimental Groups

It was experimentally hypothesized that the obtained first essay total scores of the experimental group would be higher than those of the control group. Although the scores between the two groups on the first essay were not statistically different, the average score of the experimental group was somewhat lower ($10.3 < 10.5$) than that of the control group, the opposite of what was predicted. The experimental hypothesis that the obtained revised total essay scores of the experimental group would be higher than those of the control group was also not supported. Again, the scores of the two groups were not statistically different; the mean control group score being 11.1

and the mean experimental group score 11.0.

In addition it was hypothesized that the number of words, syllables per word, and token-type ratio of the experimental group would be higher than that of the control group. Again, no statistically significant differences were found. However, the means did indicate a trend in the predicted direction. The mean number of words used by the experimental group on the first essay was 351 compared to 341 for the control group, the mean syllables per word used by the experimental group were 1.60 compared to 1.57 for the control group, and the mean token-type ratio was .376 compared to .363.

Upon revision, the lack of statistically significant differences between the two groups prevailed. The number of words used by the experimental group, however, was significantly different, but in the opposite direction predicted; the experimental group wrote fewer words ($\bar{X} = 299$) as compared to the control group ($\bar{X} = 314$). The two groups showed no significant differences in syllables per word used, although the means were in the predicted direction, 1.60 and 1.54, respectively, for the experimental and control groups. Finally, the two groups showed no statistically significant differences with regard to the token-type ratio of the revised essays; the means were in the predicted direction, with the experimental group producing a .385 ratio and the control group, .336.

Differences Between First and Revised Essays

It was experimentally hypothesized that the obtained essay total scores would reflect improvement in writing ability over time within both the control and experimental groups. An interaction was

not predicted. The experimental group was expected to show higher scores than the control group. Overall, revising the essays did not result in higher scores, $F(1,29) = 2.02$, $p > .05$. Subsequent t -tests revealed that the experimental group did not improve over time, $t(14) = -1.20$, $p > .05$. The same held true for the control group, $t(15) = -0.86$, $p > .05$.

There was a statistical effect of revision, however, $F(1,29) = 5.84$, $p < .003$. Specifically, the difference could be found in the number of words written by the experimental group, $t(14) = 3.67$, $p < .003$. The experimental group students wrote an average of 351 words on the first essay compared to an average of 299 on the revised version. Although the control group wrote less words ($\bar{X} = 314$) on the revised essay compared to the first ($\bar{X} = 341$), the difference was not significant. Neither the syllables per word nor the token-type ratio were affected by revision.

Concurrent Validity and Construct Validity

Concurrent validity was assessed by correlating grades in Writing I with grades in General Psychology and by correlating those grades with assigned essay scores (both first and revised).

Construct validity was explored by correlating assigned essay scores (both first and revised) with their corresponding number of words, syllables per word, and token-type ratio.

Assessing both the concurrent and construct validity of this study at this point is an academic exercise, as both methodologies failed to achieve reliability. The Pearson product moment correlation coefficients ranged from $-.49$ to $.47$. The correlations are

inconsistent: 13 of the 24 correlations are negative, and only one is statistically significant in the predicted direction. Several different conclusions could be drawn here: 1) The micro-elements assessed here do not assess the same qualitative characteristics of the essays as does the holistic essay scoring; 2) Neither the measurements of the micro-elements nor the holistic scores are reliable for this sample; or 3) Use of these two methods was inappropriate for this type of assessment. Table 7 shows the correlations.

Summary of Findings

It was predicted in this study that the experimental group, which received explicit cues in their instructions for a writing assignment, would perform better than the control group which did not receive the explicit instructions. It was also predicted that upon revision of the essays both groups would see improvement in their essays. Neither of these predictions was confirmed. In fact, the mean score of the experimental group on the first essay (10.3) was somewhat lower than the corresponding mean score of the control group (10.5). The mean scores of both groups showed some improvement upon revision, but not enough to be statistically significant. Again, the experimental group continued to show a lower mean score (11.0 < 11.1).

Use of the holistic scoring method failed to produce a normal distribution of scores in this study. There was also an extremely low level of intra- and interreader reliability.

Assessment of the control and experimental groups with quantification of micro-elements failed to detect any differences on the first essay. Statistical analysis of the micro-elements did detect

Table 7

Pearson Product Moment Correlations Between Writing I and
General Psychology and Micro-Elements (Number of Words,
Syllables per Word, and Token-Type Ratio) of First and
Revised Essays

| | Writing I | General Psychology | First Essay Score | Revised Essay Score |
|-----------------------|-----------|-----------------------|----------------------|------------------------|
| General Psychology | -.19 | | | |
| First Essay | | | | |
| Score | .02 | .20 | | |
| Number of Words | -.08 | .38 | .39 | |
| Syllables per Word | -.11 | .36 | -.32 | |
| Token-Type Ratio | -.16 | -.18 | -.49* | |
| Revised Essay | | | | |
| Score | -.00 | .29 | .47* | |
| Number of Words | -.22 | .10 | | -.10 |
| Syllables per Word | -.13 | .16 | | .12 |
| Token-Type Ratio | -.36 | .02 | | -.13 |

* $p < .01$

a difference in the two groups upon revision, specifically in the number of words written. The real difference was to be found in a statistically significant decrease in the mean number of words used by the experimental group in the revised essay (299 words in the revised essay, 351 words in the first essay).

The student failed to achieve concurrent validity, the assigned scores did not show a relationship to grades assigned to the students in their Writing I and General Psychology classes. Assuming grades in Writing I and General Psychology differentiate students of varying abilities, it must be concluded that both methods were unable to adequately differentiate the writing abilities of the students in this study. Nor were the methods capable of detecting improvement over time. Comparisons between the two methodologies, holistic essay scoring and quantification of micro-elements, failed to demonstrate construct validity. Individually, the two methods were unreliable, and the relationship between the two was inconsistent and weak.

The next chapter of this study will attempt to interpret and explain these findings in light of the earlier literature review.

CHAPTER V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

Summary of Design

The purpose of this study is to determine whether intervention in the classroom at the college level can aid students in improving their writing. Thirty-one students enrolled in a General Psychology class at Loyola University of Chicago were split into two groups, an experimental group ($N = 15$) and a control group ($N = 16$). Initially the control group was requested to write a one-page review or critique of a journal article, while the experimental group received the same assignment with five additional cues, as proposed by Black (1982). It was proposed that these additional cues would be used by the students at the process level of writing. These cues were not intended to force the students to develop entirely new writing processes, but to serve as prompts within their preexisting internalized writing schemas. The essays written in response to these assignments determined the initial assessment of writing abilities. With this in mind, it was hypothesized that the experimental group would write better essays than the control group when assessed with the holistic scoring method. Differences in frequencies of micro-elements were also predicted.

Four weeks later, both groups of students were requested to

revise or rewrite their essays. Both groups were given the same instructions. This second set of essays determined the final assessment of writing abilities. It was hypothesized that both groups would produce improved essays after revising or rewriting and it was not expected that one group would improve more than the other. Again, the holistic essay scoring method was used to assess the essays, and micro-elements were also counted.

In addition to scoring the essays with the holistic essay scoring method, the frequencies of several micro-elements (number of words, syllables per word, and token-type ratio) were also used in determining qualitative differences between the groups both before and after revision. The length of the essay, as measured by the number of words, served as an indication of the invention skills of the writer (Faigley et al., 1981). As proposed by Hörmann (1979) the token-type ratio (number of different words in the text over total number of words) assessed lexical restriction. Syllables per word were compared to evaluate language complexity. Use of two methodologies to assess the first and revised essays permits a comparison not only of groups and instructional intervention, but also of the methodologies themselves.

The data collected for this study included first and revised essay scores assigned by independent readers, number of words, syllables per word and token-type ratio. Intrareader reliability was assessed by correlating scores assigned to the same essays at different points in time, while interreader reliability was assessed by correlating scores between the two readers (scorers). Concurrent

validity was determined by correlating total essay scores with grades in Writing I and General Psychology. Construct validity was determined by correlating the micro-elements with assigned scores.

Findings

All of the data analyses were conducted with the SPSSx (1983) statistical package. Scores between and within groups were compared with analysis of variance tests. Micro-element differences between groups were assessed with a doubly multivariate analysis of variance test. Concurrent and construct validity as well as intra- and inter-reader reliabilities were assessed with Pearson product moment correlations. Descriptive statistics included means, standard deviations, and Pearson product moment correlations. A .01 level of significance was required so as to reduce the probability of spurious statistically significant findings.

A one-way analysis of variance failed to show support for the experimental hypothesis that the control and experimental groups differed with respect to scores obtained on the first essay. The hypothesis that the groups differed with respect to scores obtained on the revised essay was also not supported. In addition, it was experimentally hypothesized that the obtained essay scores would reflect improvement in writing ability over time within both the control and experimental groups. Again, analysis of variance failed to detect any differences between the first and revised essays of either group.

A doubly multivariate analysis of variance was conducted to compare differences between the groups' use of micro-elements.

The hypothesis that the experimental group would use more words, use more syllables per word, and produce a higher token-type ratio than the control group on the first essay was not confirmed. Upon revision the experimental group did significantly decrease the number of words they used from the number used in the first essay as well as compared to the control group. Neither the control group nor the experimental group showed any changes in syllable per word or token-type ratio after revision.

It may have been unrealistic to expect statistically significant differences in frequencies of micro-elements between groups and over time, particularly with respect to syllables per word and token-type ratio. The variability of all the dependent measures (with the exception of number of words) was exceptionally low, possibly indicating a homogeneous group. A homogeneous group, presenting little variability, will not allow for a normal bell-shaped distribution of the dependent variables. Hence, the dependent measures may vary as a function of some other factor besides general writing ability. Number of words, syllables per word and token-type ratio may therefore not be appropriate discriminators of writing ability when groups are relatively homogeneous. The decrease in number of words written lends support to research indicating that students feel they need to be more concise in their writing (Sommers, 1980). Revision may simply have provided these students with an opportunity to be more concise, but they failed to recognize the need to improve their essays' content and organization. The lack of improvement in the groups' essays over time provides supporting evidence for Stallard's

(1972) proposition that student writers demonstrate a decided lack of concern for a previously determined structure of paragraphs or the content of their texts upon revision.

Use of the holistic essay scoring method requires nearly perfect intra- and interreader reliability (Diederich, 1966). This study failed to achieve adequate levels of intra- and interreader reliability. However, these low reliabilities may be a function of the group's homogeneity. Perhaps the readers felt that the essays were qualitatively equivalent, but were obligated to assign different scores. Low reliability in combination with low concurrent and low construct validity will severely limit interpretation of statistical results.

An interesting finding, not predicted at the start of the study, but intuitively obvious, was the high correlation between scores on the first essays and their revised counterparts ($r = .47$, $p < .01$). Of all the dependent measures assessed in this study, the only consistent factor to be found was the students' own writing. This finding lends credence to the findings of Perl (1979) that writers have consistent and internalized writing processes. The lack of change between groups and over time may be attributable to the experimental assignment's inability to penetrate and modify these processes.

In addition to these statistical analyses with assigned scores and micro-elements, the students completed a brief questionnaire which may add further clarification. The students were asked to rate their own writing on a scale of 1 to 7 (1-low, 4-average, 7-

high). Of the 26 students who responded, seven rated themselves as a 6 on the scale, twelve rated themselves 5, four rated themselves 4, two rated themselves 3, and one rated him/herself a 2. In other words, 23 of the 29 (79.31%) students considered themselves to be average or better writers, and 19 of the 29 (65.52%) rated themselves to be better than average writers.

The students were also asked to write down what they thought was the one best thing that they could do to improve their writing. Five students thought they should be more careful in grammar and mechanics, seven thought they should spend more time writing, three wanted to use better vocabulary, three would like to expand rather than reduce their thoughts, two students would like to develop better style, and two would like to be more specific in their examples. One student wanted to organize thoughts better, one to read more, one to read compositions aloud to a friend for an opinion, one to practice sentence structure, one to improve spelling, and one to be neater. Only four of the students' fourteen suggestions for improvement focus on writing processes (organize thoughts better, develop better style, read aloud to a friend, and expand rather than reduce thoughts) rather than traditional writing problems of grammar and mechanics.

It thus appears that the students in this sample do not consider their writing processes to be a problem nor do they consider themselves to be inadequate writers.

Conclusions

It is nearly impossible to determine if the additional

instruction of cues and revision assisted students in writing better essays. Based on average scores, one would have to conclude that the general statement "write a review or critique" was just as effective as providing students with five additional cues as suggested by research in composition (specifically, Black [1982]). Requesting a revision of the essays appears not to have prompted the students to reevaluate and improve upon their essays. Perhaps the students were able to produce essays of improved quality but the readers and micro-elements used in this study were incapable of detecting the differences. The assignment itself may have been inappropriate; requesting a review or critique might not have allowed the students to clearly express their writing skills. The journal article itself may not have provided a conducive subject for these students to respond to and in turn may have hindered their writing performance.

Future research assessing improvement between groups and/or over time will need to provide more stringent scoring criteria and permit greater variance among the students so as to allow for acceptable reliability and sensitivity. However, the methodology used in the current study points to some real problems in college writing assessment today, especially in classes outside of the English department.

Writing across the curriculum encourages teachers of all disciplines to incorporate and evaluate student writing. There is no reason to expect instructors in disciplines outside of English to be trained any better in writing assessment than the graduate students used in this study to score essays. If this is indeed the

case, researchers and educators need to be concerned with how staff of departments other than English can assist and evaluate student writing. This is not to say that students do not need help with their writing, but that the division of labor and specialization have rendered some individuals unprepared to do so.

Little research has been conducted to determine whether training educators working outside the English department is effective. Workshops in writing across the curriculum are being held nationwide, but their short-term and long-term effectiveness has yet to be assessed. The methods being promoted by the literature have not been subjected to empirical verification.

This study looked at only a small piece of the proposals set forth by writing across the curriculum and the lack of statistically significant results does not merit concluding that its ideas are worthless. As was stated earlier, true assessment of writing across the curriculum would necessitate following students throughout their college careers and carefully studying transfer and improvement in all areas of the students' work. It would be unfair to college students to deny them opportunities to improve their writing simply because researchers do not have adequate assessment tools. Nevertheless, researchers need to work with educators in developing hypotheses as to why and how some instructional activities work better than others, why some students write better than others, and why some instructional activities help some students more than others. If, indeed, there is a crisis in student writing today, researchers and educators cannot continue to promise improvement with questionable

techniques.

Implications for Education

The critical result of this study impacting on education is the fact that there are no easy solutions to the writing problem. A writing program implemented haphazardly will do little to improve students' writing skills, and will waste valuable instructor time. Instructors outside of the English department generally are not trained to teach students to write or to evaluate student writing (Gelb, 1982). It is not clear how effective instructor training is, or whether learned procedures are implemented in the classroom as intended.

Walvoord (1972) has provided four indices of good writing: 1) good writing clearly indicates the relative importance of ideas, 2) good writing accurately reflects the relation among ideas, 3) good writing is economical, and 4) good writing is concrete, precise, simple, and vivid. By the time a student reaches college level academics, he/she recognizes the need for these writing skills, but previously internalized writing processes (Perl, 1979) may prevent relearning how to write. Teaching students good writing skills may best be implemented at earlier education levels (Hendrix, 1981) when cognitive schemas are still developing and malleable. It may only be possible for the college level instructor to fine-tune student writing skills.

Educators incorporating new programs in their writing curricula will need to work with researchers so that effectiveness and utility can be determined. Implementing new programs will require the

cooperation of educators and researchers, but the students too must recognize that these programs are being designed for their benefit. The students in this study, overall, did not feel that their writing was inadequate. Individuals not recognizing the existence of a problem, in this case poor writing, may not be ready to develop and utilize new skills and to accept assistance from educators.

Implications for Research

The results of this study are confounded by several factors. It is impossible to state confidently that the addition of cues in the assignment of the experimental group was ineffective. The two methods used for assessment, holistic essay scoring and frequency of micro-elements, proved to be unreliable and lacked validity. For the same reasons it cannot be confidently concluded that revision did not result in improvement in the essays. In tightly controlled situations, holistic essay scoring has proven itself to be a reliable and valid assessment technique (Powills, 1979). In this study, the reading and scoring session was controlled; but still, the readers tired and the essay topic was abstract. The end product was inconsistent essay scores. However, the scoring session was realistic: College educators request essays on abstract topics and they tire while grading papers. Hence, establishing the existence of improvement in student writing in the real-life college environment may prove to be a difficult task. In combination with the homogeneity frequently found in college classrooms, resulting in low variability, current assessment methods may prove to be inadequate.

A review of existing research in student writing has shown that

it is not so much that students do not know the rules and mechanics of writing: but that they lack style, ability to revise and the ability to identify and write for a particular audience (Glatthorn, 1982; Stallard, 1972; Walvoord, 1982). The cues incorporated into the experimental group assignment in this study were designed to facilitate student writing in these weak areas. Apparently, they were ineffective. Future research needs to develop hypotheses as to what sorts of cues will indeed aid students in these areas. At the same time, there appears to be a need for college level curricula to incorporate classes in the less objective areas of writing processes, such as content and organization (Freedman, 1979). As was stated earlier, by the time a student has entered college, writing processes have become internalized and consistent. Future research needs to consider when in the course of cognitive development this occurs. Intervention at the time of writing process development may prove capable of preventing the internalization of inadequate and/or inappropriate writing skills. Researchers will be doing educators a service by proposing new teaching techniques and exercises, piloting, assessing, and modifying these proposals.

REFERENCES

- Bamberg, B. (1978). Composition instruction does make a difference: A comparison of the high school preparation of college freshmen in regular and remedial English classes. Research in the Teaching of English, 12, 47-59.
- Bereiter, C., & Scardamalia, M. (1983). Levels of inquiry in writing research. In P. Mosenthal, L. Tamor, & S. A. Walmsley (Eds.), Research on writing (pp. 3-25). New York: Longman.
- Black, J. B. (1982). Psycholinguistic processes in writing. In S. Rosenberg (Ed.), Handbook of applied psycholinguistics (pp. 199-216). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.
- Bracewell, R. J. (1983). Investigating the control of writing skills. In P. Mosenthal, L. Tamor, & S. A. Walmsley (Eds.), Research on writing (pp. 177-203). New York: Longman.
- Brown, F. G. (1983). Principles of educational and psychological testing. New York: Holt, Rinehart & Winston.
- Burtis, P. J., Bereiter, C., Scardamalia, M., & Tetroe, J. (1983). The development of planning in writing. In B. M. Kroll & G. Wells (Eds.), Explorations in the development of writing (pp. 153-174). New York: John Wiley & Sons.
- Coffman, W. E. (1971). On the reliability of ratings of essay examinations in English. Research in the Teaching of English, 5, 24-36.
- Cohen, A. M. (1973). Assessing college students' ability to write compositions. Research in the Teaching of English, 7, 356-371.
- College of Arts and Sciences, Loyola University of Chicago. Chronology of writing program, policies, administrative changes, p. 1.
- Cooper, C., & Matsushashi, A. (1983). A theory of writing process. In M. Martlew (Ed.), The psychology of written language (pp. 3-40). New York: John Wiley and Sons.
- Corbett, E. P. J. (1981). The status of writing in our society. In M. F. Whiteman (Ed.), Writing: The nature, development, and teaching of written communication (pp. 47-52). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.

- Daly, J. A., & Miller, M. D. (1975). Apprehension of writing as a predictor of message intensity. The Journal of Psychology, 89, 175-177.
- Daly, J. A. (1978). Writing apprehension and writing competency. Journal of Educational Research, 72, 1, 10-14.
- de Beaugrande, R. (1984). Text production. Norwood, N.J.: Ablex Publishing Corporation.
- Diederich, P. B. (1966). How to measure growth in writing ability. English Journal, 55, 435-449.
- Della-Piana, G., Odell, L., Cooper, C., & Endo, G. (1976). The writing skills decline: So what? Educational Technology, 16, 7, 30-39.
- Faigley, L., Daly, J. A., & Witte, S. P. (1981). The role of writing apprehension in writing performance and competence. Journal of Educational Research, 75, 1, 16-21.
- Flower, L. (1979). Writer-based prose: A cognitive basis for problems in writing. College English, 4, 1, 19-37.
- Freedman, S. W. (1979). How characteristics of student essays influence teachers' evaluations. Journal of Educational Psychology, 71, 3, 328-338.
- Gelb, I. J. (1982). A study of writing. Chicago: The University of Chicago Press.
- Glatthorn, A. A. (1981). Writing in the schools. Reston, VA: National Association of Secondary School Principals.
- Glynn, S. M., Britton, B. K., Muth, D. K., & Kukhet, D. (1982). Writing and revising persuasive documents: Cognitive demands. Journal of Educational Psychology, 74, 4, 557-567.
- Gregg, N. (1983). College learning disabled writer: Error patterns and instructional alternatives. Journal of Learning Disabilities, 16, 6, 334-338.
- Hays, J. R., & Flower, L. S. (1983). Uncovering cognitive processes in writing: An introduction to protocol analysis. In P. Mosenthal, L. Tamor, & S. A. Walmsley (Eds.), Research on writing (pp. 207-220). New York: Longman.
- Heath, S. B. (1981). Toward an ethnohistory of writing in American education. In M. F. Whiteman (Ed.), Writing: The nature, development, and teaching of written communication (pp. 25-45). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.

- Hendrix, R. (1981). The status and politics of writing instruction. In M. F. Whiteman (Ed.), The nature, development, and teaching of written communication (pp. 53-70). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.
- Hörmann, H. (1976). To mean - To understand. New York: Springer-Verlag.
- Hörmann, H. (1979). Psycholinguistics. New York: Springer-Verlag.
- Jerabek, R., & Dieterich, D. (1975). Composition evaluation: The state of the art. College Composition and Communication, 26, 183-186.
- Kidder, L. H. (1981). Sellits, Wrightsman & Cooks research methods in social relations. New York: Holt, Rinehart, & Winston.
- Klein, S. P., & Hart, F. M. (1968). Chance and systematic factors affecting essay grades. Journal of Educational Measurement, 5, 3, 197-206.
- Kowal, S., & O'Connell, D. C. (in press). Writing as language behavior. In A. Matsushashi (Ed.), Writing in real times: Modelling production processes. New York: Longman.
- Mackay, A. L. (1984). In the matter of man. Science, 24, 1, 13-14.
- Maimon, E. P., Belcher, G. L., Hearn, G. W., Nodine, B. F., & O'Connor, F. W. (1981). Writing in the Arts and Sciences. Cambridge, MA: Winthrop Publishers, Inc.
- Martlew, M. (Ed). (1983). The psychology of written language. New York: John Wiley and Sons.
- Mosenthal, P. (1983). On defining writing and classroom writing competence. In P. Mosenthal, L. Tamor, & S. A. Walmsley (Eds.), Research on writing (pp. 26-71). New York: Longman.
- Murray, D. M. (1978). Internal revision: A process of discovery. In C. Cooper & L. Odell (Eds.), Research on composing (pp. 85-104). Urbana, IL: National Council of Teachers of English.
- National Institute of Education (1984). Planning grants and institutional grants for educational research and development centers. Washington, D.C.: U.S. Department of Education.
- Odell, L., Cooper, C. R., & Courts, C. (1978). Discourse theory: Implications for research in composing. In C. Cooper & L. Odell (Eds.), Research on composing: Points of departure (pp. 1-12). Urbana, IL: National Council of Teachers of English.

- Odell, L. (1981). Defining and assessing competence in writing. In C. Cooper (Ed.), The nature and measurement of competency in English (pp. 95-138). Urbana, IL: National Council of Teachers of English.
- Olson, D. R., & Torrence, N. (1983). Writing and criticizing texts. In B. M. Kroll & G. Wells (Eds.), Explorations in the development of writing (pp. 31-42). New York: John Wiley and Sons.
- Osgood, C. E., & Sebeok, T. A. (Eds.) (1965). Psycholinguistics. Bloomington, IN: Indiana University Press.
- Perl, S. (1979). The composing processes of unskilled college writers. Research in the Teaching of English, 18, 4, 317-336.
- Pianko, S. (1979). A description of the composing process of college freshman writers. Research in the teaching of English, 13, 5-22.
- Powills, J. A. (1979). Holistic essay scoring: An application of the model for the evaluation of writing ability and the measurement of growth in writing ability over time. Master's Thesis, Loyola University of Chicago.
- Rosenberg, S. (Ed.). (1982). Handbook of applied psycholinguistics. Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.
- Slobin, D. I. (1979). Psycholinguistics. Palo Alto, CA: Scott, Foresman & Company.
- Sommers, N. (1980). Revision strategies of student writers and experienced adult writers. College Compositions and Communication, 31, 378-387.
- SPSSx Inc. (1983). SPSSx user's guide. Chicago.
- Stallard, C. K. (1972). An analysis of the writing behavior of good student writers. Research in the Teaching of English, 8, 206-218.
- Walvoord, B. E. (1982). Helping students write well: A guide for teachers in all disciplines. New York: The Modern Language Association of America.
- Warshauer, S. F., & Calfee, R. C. (1983). Holistic assessment of writing: Experimental design and cognitive theory. In P. Mosenthal, L. Tamor, & S. A. Walmsley (Eds.), Research on writing (pp. 75-98). New York: Longman.

Weiner, E. S. (1980). Diagnostic evaluation of writing skills.
Journal of Learning Disabilities, 13, 1, 43-48.

Winer, B. J. (1971). Statistical principles in experimental design. New York: McGraw-Hill Book Company.

APPENDIX A

Sample First Essay from Experimental Group:

The French physician Julien Offroy de la Mettrie was in serious trouble for expressing his extraordinary ideas and others took after too. He wrote a book and was forced out of Paris for it. In the Netherlands he published another book, expressing his ideas about a man being a machine. Craftsmen took pride for building robots and had them on display as fine collections of androit-meaning manlike in Greek.

Continuos strand of thoughts from Leucippus and Democritus reported "Nothing exists but atoms and empty space". Pierre Simon Laplace carried atomist and reductionist arguments. Using computers idea's limitations are clearly understood but we still rely on statistical averages because the computer has to guess. Behavior of populations or atoms may go one way or another. Leo Tolstoy took a molecular dynamics approach considering a possibility to make only predictions. There was an extreme controversy between ideas about atoms, man, and approaches to these.

Machines will appear to have more free will since designs were improving. Grey Walter designed and built electric tortoise, with primitive automatons with two tropisms- meaning an innate tendency to react to a stimulus.

Indeed, there is a great complexity. The difference between the living and the mechanical will fade away. Reductionism--meaning reducing the complex data to simple terms, will vanish. Our sense of free will would be the same.

· Sample Revised Essay from Experimental Group:

The article "In The Matter Of Man" presents a number of controversial opinions. One of these is the idea that man is only a complicated machine. This idea was ascribed principally by Julien Offroy de La Mettrie in two books that he published. This strong belief is traceable all the way back to Greek thinkers.

Pierre-Simon Laplace carried atomist and reductionist arguments in the nineteenth century. Reasoning from the principles of reductionism, Laplace argued that everything in the universe is predictable. Unfortunately, this is not the case: even atoms are unpredictable in their interactions, and the domain of human activity is far more complicated and unpredictable. Some of this idea's limitations were clarified by the use of computers in experiments. Indeed, in all cases predictability must be based in statistics.

The author predicts that man will retain his conviction of free will, but that the ability of scientists to predict the behavior at the atomic level and the human level will continue to improve. The difference between the behavior of the living and the mechanical will diminish. The question of reductionism will therefore gradually vanish as we learn and experience more about the causes of our behavior.

Sample First Essay from Control Group:

In the article, "In the Matter of Man," the author discusses the beliefs of three men concerning the human soul, atoms and the ability to see into the future. The first man Julien Offroy de La Mettrie believed that man was "no more than a complicated machine." By this he feels that our so-called soul is all in the mind and when the body dies, so does the soul.

Lucretius believed in the same principle as La Mettrie. Lucretius said in the first century B.C. that man which has emotional capability is made up of atoms which have no emotional capability.

Simon Laplace, a mathematician, took the two views (atomist and reductionist) one step further. He said if there were such a formula able to predict the movements of certain objects, and was applicable to all forms of matter. We could then predict future, and reverse the formula to understand the past.

Such a formula would be impossible, for a computer would sooner or later come to a situation that is impossible to give a correct answer. The computer must therefore guess and by doing so the formula would become invalid. If the trajectory of one atom cannot be predicted with complete accuracy then the theory is disproved.

In conclusion, man does have free will, but we know that molecules in our mind must collide for any action to take place.

Sample Revised Essay from Control Group:

In the article, "In the Matter of Man," the author discusses the beliefs of three men concerning the human soul, atoms and the ability to see into the future. The first man Julien Offroy de La Mettrie believed that man was "no more than a complicated machine." By this he feels that our so-called soul is all in the mind and when the body dies, so does the soul.

Lucretius believed in the same principle as La Mettrie. Lucretius said in the first century B.C. that man (which has emotional capability) is made up of atoms which have no emotional capability.

Simon Laplace a mathematician, took the two views (atomist and reductionist) one step further. He said if such a formula existed to predict the movements of certain objects, and was applicable to all forms of matter, we could then predict the future, and reverse the formula to understand the past.

Such a formula would be impossible, for a computer would sooner or later come to a situation for which it is impossible to give a correct answer. The computer must therefore guess and by doing so the formula would become invalid. If the trajectory of one atom cannot be predicted with complete accuracy then the theory is disproved.

In conclusion, man does have free will, but we know that molecules in our mind must collide for any action to take place.

SCORE SHEET

READER NAME _____

ESSAY NUMBER _____

SCORE _____

SCORE SHEET

READER NAME _____

ESSAY NUMBER _____

SCORE _____

APPENDIX B

RAW DATA

EXPERIMENTAL GROUP (N = 15)

| Writing I | General Psychology | FIRST ESSAY | | | REVISED ESSAY | | |
|-----------|--------------------|-----------------|----------------|------------------|-----------------|----------------|------------------|
| | | Number of Words | Syllables/Word | Token-type Ratio | Number of Words | Syllables/Word | Token-type Ratio |
| 4 | 3 | 328 | 1.51 | .384 | 355 | 1.62 | .358 |
| 3 | 2 | 497 | 1.56 | .308 | 402 | 1.55 | .311 |
| | 3 | 459 | 1.55 | .285 | 352 | 1.46 | .301 |
| 4 | 2 | 302 | 1.59 | .404 | 332 | 1.52 | .419 |
| 3 | 4 | 365 | 1.54 | .301 | 285 | 1.50 | .358 |
| 3 | 2 | 394 | 1.63 | .297 | 263 | 1.76 | .384 |
| 3 | 2 | 253 | 1.54 | .486 | 294 | 1.51 | .333 |
| 3 | 2 | 259 | 1.49 | .340 | 230 | 1.53 | .365 |
| 3 | 4 | 468 | 1.73 | .323 | 360 | 1.67 | .367 |
| | 1 | 225 | 1.73 | .547 | 198 | 1.78 | .475 |
| 2 | 4 | 363 | 1.52 | .375 | 321 | 1.64 | .364 |
| 4 | 4 | 441 | 1.62 | .360 | 347 | 1.52 | .360 |
| 4 | 3 | 237 | 1.75 | .380 | 209 | 1.68 | .517 |
| 4 | 4 | 348 | 1.72 | .431 | 255 | 1.78 | .463 |
| 4 | 2 | 331 | 1.60 | .417 | 288 | 1.56 | .406 |

RAW DATA - EXPERIMENTAL GROUP (N = 15) (continued)

FIRST ESSAY

| Reader 1 Time 1 | Reader 1 Time 2 | Sum Reader 1 | Reader 2 Time 1 | Reader 2 Time 2 | Sum Reader 2 |
|--------------------|--------------------|-----------------|--------------------|--------------------|-----------------|
| 4 | 4 | 8 | 3 | 3 | 6 |
| 4 | 4 | 8 | 3 | 2 | 5 |
| 3 | 4 | 7 | 4 | 3 | 7 |
| 2 | 2 | 4 | 4 | 2 | 6 |
| 3 | 2 | 5 | 3 | 1 | 4 |
| 4 | 3 | 7 | 4 | 1 | 5 |
| 2 | 2 | 4 | 3 | 3 | 6 |
| 2 | 1 | 3 | 2 | 3 | 5 |
| 3 | 2 | 5 | 2 | 3 | 5 |
| 1 | 1 | 2 | 1 | 1 | 2 |
| 2 | 2 | 4 | 2 | 2 | 4 |
| 4 | 3 | 7 | 4 | 4 | 8 |
| 4 | 2 | 6 | 1 | 1 | 2 |
| 3 | 3 | 6 | 2 | 2 | 4 |
| 4 | 4 | 8 | 1 | 1 | 2 |

RAW DATA - EXPERIMENTAL GROUP (N = 15) (continued)

| Reader 1 Time 1 | Reader 1 Time 2 | Sum Reader 1 | Reader 2 Time 1 | Reader 2 Time 2 | Sum Reader 2 | Total Score Essay 1 | Total Score Essay 2 |
|--------------------|--------------------|-----------------|--------------------|--------------------|-----------------|------------------------|------------------------|
| 4 | 4 | 8 | 4 | 4 | 8 | 14 | 16 |
| 4 | 4 | 8 | 1 | 2 | 3 | 13 | 11 |
| 4 | 4 | 8 | 1 | 3 | 4 | 14 | 12 |
| 2 | 2 | 4 | 2 | 2 | 4 | 10 | 8 |
| 4 | 2 | 6 | 2 | 2 | 4 | 9 | 10 |
| 3 | 2 | 5 | 3 | 3 | 6 | 12 | 11 |
| 4 | 3 | 7 | 3 | 3 | 6 | 10 | 13 |
| 2 | 1 | 3 | 3 | 3 | 6 | 8 | 9 |
| 3 | 2 | 5 | 2 | 4 | 6 | 10 | 11 |
| 2 | 2 | 4 | 2 | 3 | 5 | 4 | 9 |
| 2 | 3 | 5 | 2 | 2 | 4 | 8 | 9 |
| 4 | 4 | 8 | 3 | 2 | 5 | 15 | 13 |
| 3 | 2 | 5 | 2 | 1 | 3 | 8 | 8 |
| 4 | 4 | 8 | 3 | 1 | 4 | 10 | 12 |
| 3 | 4 | 7 | 3 | 3 | 6 | 10 | 13 |

RAW DATA

CONTROL GROUP (N = 16)

| Writing I | General Psychology | FIRST ESSAY | | | REVISED ESSAY | | |
|-----------|--------------------|-----------------|----------------|------------------|-----------------|----------------|------------------|
| | | Number of Words | Syllables/Word | Token-type Ratio | Number of Words | Syllables/Word | Token-type Ratio |
| 3 | 1 | 275 | 1.58 | .393 | 239 | 1.65 | .372 |
| | 0 | 450 | 1.41 | .267 | 438 | 1.42 | .258 |
| 3 | 4 | 460 | 1.59 | .263 | 304 | 1.63 | .296 |
| 4 | 4 | 451 | 1.59 | .446 | 275 | 1.47 | .400 |
| 4 | 3 | 257 | 1.49 | .393 | 282 | 1.50 | .277 |
| 4 | 3 | 385 | 1.72 | .418 | 245 | 1.68 | .469 |
| 2 | 1 | 346 | 1.41 | .306 | 270 | 1.33 | .341 |
| | 4 | 338 | 1.66 | .281 | 308 | 1.56 | .243 |
| | 2 | 256 | 1.55 | .422 | 372 | 1.41 | .298 |
| 2 | 1 | 229 | 1.56 | .432 | 233 | 1.55 | .373 |
| | 3 | 328 | 1.48 | .283 | 292 | 1.61 | .315 |
| 4 | 2 | 246 | 1.73 | .411 | 225 | 1.69 | .342 |
| 3 | 2 | 478 | 1.64 | .406 | 365 | 1.62 | .449 |
| 3 | 2 | 278 | 1.58 | .450 | 314 | 1.55 | .369 |
| | 1 | 281 | 1.50 | .317 | 371 | 1.55 | .280 |
| 4 | 4 | 409 | 1.69 | .323 | 498 | 1.49 | .289 |

RAW DATA - CONTROL GROUP (N = 16) (continued)

FIRST ESSAY

| Reader 1 Time 1 | Reader 1 Time 2 | Sum Reader 1 | Reader 2 Time 1 | Reader 2 Time 2 | Sum Reader 2 |
|--------------------|--------------------|-----------------|--------------------|--------------------|-----------------|
| 2 | 2 | 4 | 3 | 3 | 6 |
| 2 | 2 | 4 | 4 | 4 | 8 |
| 3 | 3 | 6 | 4 | 3 | 7 |
| 2 | 2 | 4 | 2 | 3 | 5 |
| 4 | 4 | 8 | 4 | 3 | 7 |
| 2 | 2 | 4 | 2 | 3 | 5 |
| 4 | 3 | 7 | 3 | 2 | 5 |
| 3 | 2 | 5 | 3 | 3 | 6 |
| 3 | 3 | 6 | 3 | 2 | 5 |
| 2 | 2 | 4 | 2 | 1 | 3 |
| 4 | 2 | 6 | 3 | 3 | 6 |
| 4 | 3 | 7 | 1 | 4 | 5 |
| 2 | 3 | 5 | 2 | 2 | 4 |
| 3 | 3 | 6 | 2 | 1 | 3 |
| 3 | 2 | 5 | 2 | 1 | 3 |
| 4 | 2 | 6 | 2 | 1 | 3 |

RAW DATA - CONTROL GROUP (N = 16) (continued)

REVISED ESSAY

| Reader 1 Time 1 | Reader 1 Time 2 | Sum Reader 1 | Reader 2 Time 1 | Reader 2 Time 2 | Sum Reader 2 | Total Score Essay 1 | Total Score Essay 2 |
|--------------------|--------------------|-----------------|--------------------|--------------------|-----------------|------------------------|------------------------|
| 3 | 3 | 6 | 3 | 3 | 6 | 10 | 12 |
| 2 | 2 | 4 | 1 | 2 | 3 | 12 | 7 |
| 4 | 4 | 8 | 1 | 3 | 4 | 13 | 12 |
| 3 | 1 | 4 | 4 | 4 | 8 | 9 | 12 |
| 2 | 3 | 5 | 2 | 3 | 5 | 15 | 10 |
| 4 | 4 | 8 | 2 | 2 | 4 | 9 | 12 |
| 2 | 3 | 5 | 4 | 3 | 7 | 12 | 12 |
| 3 | 3 | 6 | 4 | 2 | 6 | 11 | 12 |
| 2 | 2 | 4 | 4 | 2 | 6 | 11 | 10 |
| 4 | 2 | 6 | 2 | 2 | 4 | 7 | 10 |
| 4 | 4 | 8 | 4 | 4 | 8 | 12 | 16 |
| 4 | 3 | 7 | 3 | 3 | 6 | 12 | 13 |
| 3 | 2 | 5 | 3 | 2 | 5 | 9 | 10 |
| 2 | 2 | 4 | 3 | 2 | 5 | 9 | 9 |
| 3 | 2 | 5 | 3 | 2 | 5 | 8 | 10 |
| 4 | 3 | 7 | 2 | 1 | 3 | 9 | 10 |

APPROVAL SHEET

The thesis submitted by Donna A. Van De Water has been read and approved by the following committee:

Rev. Daniel C. O'Connell, S.J., Director
Professor, Psychology, Loyola

Dr. Deborah J. Holmes
Professor, Psychology, Loyola

The final copies have been examined by the Director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the Committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts.

April 14, 1985
Date

Daniel C. O'Connell
Director's Signature