

To the Graduate Council:

I am submitting herewith a dissertation written by Ruixuan Mao, entitled "An Analytical Study of Twenty Research Reports in the Teaching of Writing." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education.

Mark A. Christiansen

Mark A. Christiansen,
Major Professor

We have read this dissertation
and recommend its acceptance:

Donald J. Lippert
Ted Hopp
John R. Hays
Allison Enos

Accepted for the Council:

Lew Minkel

Associate Vice Chancellor and
Dean of The Graduate School

AN ANALYTICAL STUDY OF
TWENTY RESEARCH REPORTS
IN THE TEACHING OF WRITING

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Ruixuan Mao
December 1993

To my colleagues
in China

ACKNOWLEDGEMENTS

I should like to express my appreciation to the members of my committee, who, with their respective expertise and unreserved concern, made every effort to insure that I completed my work in time to receive the Doctor of Philosophy degree in December, 1993. I extend my thanks to Mark A. Christiansen, chair; Theodore W. Hipple; John R. Ray; Donald J. Dessart; and Allison R. Ensor. I should also like to thank my colleagues in the field of teaching writing who permitted me to use their research work freely in the course of this study.

Finally, I should like to thank my wife, Mi Hu, who has made completing the dissertation and the degree all worthwhile.

ABSTRACT

The purpose of this study was a unique one. The author of this study was an English professor from the People's Republic of China. Although more than thirty million people in China had been exposed to some degree of English education, research in the teaching of writing in English was practically nonexistent. Furthermore, it was the author's desire ever since he came to the United States for graduate studies that the expertise acquired in the area of writing instruction in America should be introduced to China.

To achieve this purpose, this study focused on twenty published research reports in the teaching of composition at the secondary level by means of content analysis. These published research reports were selected from either the NCTE research report series or articles published in leading journals in the area of secondary composition instruction. By doing so, the author of this study expected to describe the basic picture of this area of study for his Chinese colleagues and to facilitate a program to bring the Chinese professionals into contact with their American colleagues.

In this study, the analysis was conducted by categorizing the research reports into several groups: descriptive studies, experimental studies, and meta-analysis. In each group, the theoretical framework, the

assumptions, the processes, and the strengths and weaknesses were discussed, involving many methodological issues and concerns.

Based on the analysis, a theory of hierarchy was suggested as an overriding theoretical structure for the research of composition. The theory of hierarchy was known to be effective in the study of complex situations. It has been proved to be successful in the study of biology, chemistry, ecology, linguistics, and many other areas of study. In psychology, the theory of hierarchy has also been found to be closely related to how the human mind works.

When the hierarchical model was adopted, various types of hierarchical structures could be explored and mathematical models could be devised to study this theory and consequently to further our understanding of the process of writing and the process of teaching composition.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Rationale for the Study	1
Selection of Reports for Analysis	4
Purpose and Design of the study	19
Definitions Used in This Study	22
II. THE HISTORY OF RESEARCH IN TEACHING WRITING	26
Early Research in Teaching Writing	26
Methodology and Foci of Research	34
Research in Teaching Writing in the Field of English as a Second Language (ESL)	37
Toward the Hierarchy Theory	44
III. EXPERIMENTAL RESEARCH IN TEACHING WRITING	51
Philosophies, Theories, and Hypotheses	53
Designs, Samples, and Procedures	66
Instruments, Statistics, and Conclusions	82
Strengths, Weaknesses, and Applicability	109
IV. DESCRIPTIVE RESEARCH IN TEACHING WRITING	114
Philosophies, Theories, and Concepts	114
Survey, Observation, and Case Studies	117
Assessment, Measurement, and Evaluation Studies	139
Strengths, Weaknesses, and Applicability	165
V. META-ANALYSIS	171
An Overview of Meta-analysis	171
Literature Reviews	176
Research Critiques and Reflective Study	192
Hillocks Jr.'s Study and the Meta-analysis	197
VI. DISCUSSION AND CONCLUSION	204
BIBLIOGRAPHY	213
APPENDICES	252
APPENDIX A	253
APPENDIX B	266
APPENDIX C	269
APPENDIX D	273
VITA	274

LIST OF TABLES

TABLE	PAGE
1. List of Coded Reports for Analysis	14
2. Design of Experiment	69
3. Score Sheet	89
4. Criteria for Maturing Style	93
5. Validation Data for Test C	93
6. Reliability of the Criteria	94
7. Rater Reliability	95
8. Four Types of Inquiry	116
9. Uses of School Writing	126
10. Dimensions of the Composing Process Among Twelfth-grade Writers	132
11. Model of a Curriculum System of Meaning	139
12. Mean Score on Three Factors of Syntactic Maturity--O'Donnel, 1967; and Hunt, 1965 and 1970	146
13. Mean Scores on Three Factors on Syntactic Maturity--Grade Ten Through Sixth Year University	147
14. Predictors of Grade Five Holistic Quality Score, n=45	152
15. Predictors of Grade Eight Holistic Quality Score, n=50	153
16. Predictors of Grade Eleven Holistic Quality Score, n=49	154
17. Halliday and Hasan Cohesion Categories	157
18. NAEP Scoring Guide: Primary-trait Coherence (Cohesion)	158
19. Best Possible Subset from Regression on the Primary-trait	160
20. Correlations of Prior Knowledge and Writing Measures	164
21. Traditional Literature Reviews	177
22. Studies of Writing/Reading Relationships	184
23. Dimensions in Writing Across the Curriculum Research	192
24. Reflective Studies	194
25. Difference Between Original Data and Modified Data in Reed, Burton, and Vandett's Study	197
26. Mean Scores on Student Essay for Holistic Quality and Words per T-unit, by Experimental Conditions of Ability, External Storage, and Heuristic	198
27. Means and Standard Deviation (in Parentheses) of Holistic Quality Scores and Words per Clause for Topic-type (Group or Individual)	199
28. Results of Meta-analysis by Different Dimensions	203

CHAPTER I

INTRODUCTION

Rationale of the Study

The incentive that invoked this study was threefold. The author of this study, a professor of English in China, had observed that research in the teaching of English composition in the People's Republic of China was still a vast unexplored territory. In China, more than thirty million people had received some degree of education in English in regular schools, evening schools, television universities, developmental programs, special-purpose programs, correspondence courses, and various types of training programs. English was a regular course in most secondary schools and all institutes of higher education. Some key-note elementary schools offered English from the third grade. However, the curricula for these programs were mostly constructed around a small number of available textbooks. Furthermore, the faculty responsible for teaching English was still in need of advanced formal training in the content areas, such as reading, literature, linguistics, and speech, as well as in the various research methodologies already known to the western academic world. The teaching of English composition in these programs was even less structured. As of 1985, the national guideline for teaching composition in the University English

curriculum was quite vague so far as research was concerned (National Commission of Education of the People's Republic of China, 1985). This guideline stipulated only the minimum amount of words required of English majors and non-English majors for their writing activities in different courses at the undergraduate level during their four-year program. Nothing was mentioned about the implementation of research findings into the teaching of English composition, since there was, as a matter of fact, very little effort, and very little funding as well, dedicated to research in teaching English composition. Though research was hindered (Liu, 1988), Chinese professionals had already begun to realize the importance of basing the teaching of English composition on significant research studies (Hou, 1987).

Research in the teaching of composition in the United States did not have a long history, either. However, significant progress had already been made in this field of study. Especially during the second half of this century, a large amount of funding and effort had been invested to explore this field. In their 1963 report, Braddock, Lloyd-Jones, and Schoer still used the metaphor of alchemy to describe this effort (Braddock, Lloyd-Jones, and Schoer, 5). Twenty-three years later, in another report by Hillocks, Jr., in 1986, Lloyd-Jones pointed out that the changes in the research on composition over the last two decades were striking (Lloyd-Jones, in Hillocks, Jr., 1986, p. xiii).

The amount and quality of research done in the United States in the teaching of composition should inspire new research effort among Chinese professionals.

Secondly, the author of this study found, after reading literature in this field, that researchers were working in various directions and along different lines, some trying to find ways that would improve the classroom teaching of composition, and some trying to grapple with the theoretical framework of composition research. A variety of models had been adopted and a variety of theories had been formulated, too. To obtain a comprehensive view of the whole field of teaching composition would require years and years of effort. The author felt that it was a feasible idea to study closely a reasonably large number of selected research reports in the teaching of composition to generate a set of guidelines for Chinese professionals in their effort to promote a better understanding both of the theoretical aspect and of the practical aspect of teaching English composition.

Thirdly, research methodology had always been a focus of interest for the author of this study. Not only would the conclusions drawn in the various research reports enhance the understanding of the teaching of English composition as a profession, but the different approaches adopted in these research projects would also offer an excellent opportunity for the author and his colleagues in

China to examine the methodological issues involved in the research itself. As a matter of fact, the methodological concerns were more important for the purpose of this study.

When the above considerations were put together, the author decided to conduct a study which could result in a comprehensive description of the research effort made in this field in the United States and a set of important guidelines for Chinese professionals to apply to their future research in the teaching of English composition in China. To achieve this purpose, the author of this study believed that the content and result of this study should:

- (1) be comprehensive enough to cover the basic aspects of research in English composition,
- (2) be informative enough to cover the basic facts about the development of the research of English composition in the United States,
- (3) applicable enough for Chinese researchers of the teaching of English composition to readily use these guidelines in their studies, and
- (4) flexible enough to benefit the teachers of English composition in their classroom practice in China.

Selection of Reports for Analysis

Based upon the above rationale, the selection of research reports for this study had to observe the following

considerations:

- (1) The reports selected for study had to be from a relatively long period to offer a certain amount of historical perspective.
- (2) The reports selected for study had to be related to various aspects of the teaching of composition.
- (3) The reports selected for study had to be by as many different researchers as possible.
- (4) The methodologies involved in the reports selected for study had to display as much variety of approaches as possible.
- (5) The reports selected for study had to possess a certain length for more detailed analysis.
- (6) The reports selected had to show the major sources from which these reports and other reports were accessible to Chinese researchers.
- (7) A number of highly recognized reports had to be included in this study for their importance in the development of research in this area.
- (8) The total number of reports for this study had to stay within a manageable limit, for example, 20.
- (9) The reports to be selected could include some from other countries where English composition was taught and studied.

Since a randomized selection would miss a number of very important research reports published, the author

decided to handpick rather than randomize, as predictive validity derived from the process of randomization in the selection of subjects for study was not among the first priorities of this study. This amount of arbitrariness could be offset by following a carefully designed set of criteria for selection for the purpose of this study.

With this understanding, the author looked into the published research reports sponsored by the National Council of Teachers of English and the American Educational Research Association, and reports carried in the following journals:

American Journal of Education

Curriculum Studies

English Journal

Journal of Education

Journal of Educational Research

Journal of Research and Development in Education

Journal of Teaching Writing

Language Arts

Research in the Teaching of English, and

ERIC Clearinghouse publications.

These organizations and journals were recognized in the academic world for their emphasis on research.

Dissertations were also an important source of information for this study. As only published research reports were used for analysis in this study, dissertations were, therefore, used for reference purposes, the reason being

that Chinese researchers' future access to these dissertations could be understandably limited, and that some of the research reports selected for this study were condensed reports of the original dissertations.

Other major limitations used for screening the reports to be included in this study were regarding the level at which the composition for research was taught, the time period in which the research studies were carried out, and the field of expertise with which these research projects were concerned. The first limitation was on the level at which research was conducted in teaching composition. Obviously, it would have been almost impossible to cover this field from kindergarten through college. The second limitation concerned the time variable. It would also have been an impossible effort to trace the history of the research in teaching composition to its very roots. The third limitation was on the specialty area. Besides the research in teaching English composition, there was a large number of research fields, such as teaching English as a second language, second language acquisition, teaching composition in a foreign language, rhetoric and composition, all branches of linguistics, discourse theory, cognitive theory, speech communication, speech psychology, and some others. Therefore, when the decision for the author was actually made regarding the scope and the time limit as well as the specialty area limit embedded in this study, the

author felt obligated to explain why he had made his selection in this study.

The limitation of level was on the secondary school. The limitation of time was on the past thirty years. The limitation of specialty area was on the teaching of secondary composition. The reasons behind these limitations were as follows:

When determining what research studies to include, the author of this study considered the needs of the secondary school students and college students who were studying English as a foreign language in China. Besides the assumed equivalency of educational background and developmental processes, the students of both countries had similar frequency of writing assignments, about once a week to once every two weeks. The length of each assignment was about one hundred words to five hundred words. The topics were chiefly related to things familiar to them. The organization and mechanics involved in these assignments were of an accumulative and developmental nature as could largely be seen in an average American high school. Therefore, it seemed most appropriate to the author of this study to make an analytical study of the research reports in teaching composition at the secondary level if this study was meant to be beneficial to Chinese students learning English composition. As some reports that were actually selected also involved mature writers, college students, or

elementary students, the author of this study decided to concentrate on the content that was directly related to secondary composition, and skip the unrelated parts unless they were significantly related to the major issues of concern as defined in the purposes for this study mentioned before.

When the limitation of time was determined, the author of this study took into consideration two aspects. One aspect concerned the time period in which publications were searched for general information, and the other aspect concerned the specific period in which the research reports should be selected for analysis. The author of this study wished to extend the time period for general information backward to the beginning of this century and to limit the time period for the selection of specific published research reports for analysis to the past thirty years. Therefore, the focus of interest was on the past thirty years, but a knowledge of what preceded 1963 would throw light on the present study.

When the limitation of specialty area was determined, the author of this study decided to use the research reports in teaching secondary English composition. The other areas might be investigated in other studies with a better focus.

Still, tens of thousands of published research reports qualified. The author of this study decided to limit the list based on a number of additional criteria. First of

all, the author picked a few very important research reports sponsored and published by the National Council of Teachers of English, which included Applebee's report; Braddock, Lloyd-Jones, and Schoer's report; Emig's report; Mellon's report; and O'Hare's report. These were the full-length reports related to secondary composition studies. Then, the author looked into research reports published in the journals listed above, making sure that the other fifteen reports were at least eight pages long, distributed more or less evenly throughout the thirty years, and that little or no repetition occurred in terms of topic or author. The considerations for these criteria were (1) to exclude shorter articles, (2) to cover as many areas and topics as possible in a limited selection, and (3) to achieve a chronological perspective in the selection of reports for analysis in this study. In the actual selection, two reports by Canadian researchers were included, since these reports were very closely related to the research efforts made by American scholars, as could be seen in the analysis of this study, and were published in an American journal.

When these criteria were carefully exercised, a list of twenty published research reports was generated as follows:

Applebee, A. N. (1981). Writing in the secondary school: English and content areas. (Report No. 21). Urbana, IL: National Council of Teachers of English.

- Beach, R. (1979). "The effects of between-draft teacher evaluation versus student self-evaluation on high school students' revising of rough drafts." Research in the Teaching of English, 13, 111-119.
- Braddock, R., Lloyd-Jones, R., & Schoer, L. (1963). Research in written composition. Urbana, IL: National Council of Teachers of English.
- Chesky, J., & Hiebert, E. H. (1987). "The effects of prior knowledge and audience on high school students' writing." The Journal of Educational Research, 80, 304-313.
- Cooper, C. (1975). "Measuring growth in writing." English Journal, 64(3), 111-120.
- Donlan, D. (1974). "Teaching writing in the content areas: Eleven hypotheses from a teacher survey." Research in the Teaching of English, 8, 250-262.
- Emig, J. (1971). The composing process of twelfth graders. (Report No. 13). Urbana, IL: National Council of Teachers of English.
- Graves, D. H., & Murray, D. M. (1980). "Revision: In the writer's workshop and in the classroom." Journal of Education, 162, 38-56.
- Grobe, C. (1981). "Syntactic maturity, mechanics, and vocabulary as predictors of quality ratings." Research in the Teaching of English, 15, 75-85.

- Hillocks, G., Jr. (1984). "What works in teaching composition: A meta-analysis of experimental treatment studies." American Journal of Education, 93, 133-170.
- Lutkus, A. (1987). "Problems in measuring syntactic development: T-units versus sentence weight." Journal of Teaching Writing, 6, 49-67.
- McCann, T. M. (1989). "Student argumentative writing knowledge and ability at three grade levels." Research in the Teaching of English, 23, 62-76.
- McCulley, G. A. (1985). "Writing quality, coherence, and cohesion." Research in the Teaching of English, 19, 269-282.
- Mellon, J. C. (1969). Transformational sentence-combining: A method for enhancing the development of syntactic fluency in English composition. (Report No. 10). Urbana, IL: National Council of Teachers of English.
- O'Hare, F. (1973). Sentence combining: Improving student writing without formal grammar instruction. (Report No. 15). Urbana, IL: National Council of Teachers of English.
- Reed, W. M., Burton, J. K., & Vandett, N. M. (1988). "Daly and Miller's Writing Apprehension Test and Hunt's T-unit analysis: Two measurement precautions in writing research." Journal of Research and Development in Education, 21(2), 1-8.

- Söter, A. O. (1987). "Recent research in writing: Implications for writing across the curriculum." Curriculum Studies, 19, 425-438.
- Steward, M. F. (1978). "Syntactic maturity from high school to university: A first look." Research in the Teaching of English, 12, 37-46.
- Stotsky, S. (1983). "Research on reading\writing relationships: A synthesis and suggested directions." Language Arts, 60, 627-642.
- Swanson-Owens, D. (1986). "Identifying natural sources of resistance: A case study of implementing writing across the curriculum." Research in the Teaching of English, 20, 69-97.

Table 1 lists the 20 research reports in their chronological sequence with a coding system. The codes will be used in the following chapters and Appendix A.

Besides, the reports selected had the following merits to justify their inclusion in this study. Braddock, Lloyd-Jones, and Schoer's report was a famous report in the early 60's published by NCTE, which summarized the status quo of the composition research done before that time, focusing on five carefully selected experimental studies. Emig's report was also a widely cited report, published in the early 70's and included in the NCTE composition research report series, which adopted a new and powerful research methodology in composition studies. Mellon's report and O'Hare's report,

Table 1
LIST OF CODED REPORTS FOR ANALYSIS

CODE	AUTHOR	THEME	LEVEL	METHODOLOGY	NUM	PAGE	YR
RA	Braddock	Research	EJHC	Summary	--	142	63
RBE	Emig	Behavior	12	Case Study	8	151	71
RBO	O'Hare	Sentence	7	Experiment	83	109	71
RC	Donlan	Hypothesis	Facul	Survey	92	13	74
RDC	Cooper	Measurement	9-12	Summary	--	10	75
RDM	Mellon	Writing	3,7,11	Measurement	2000	127	75
REC	Stewart	T-unit	10-18	Measurement	302	10	78
RF	Beach	Revision	10-12	Experiment	103	9	79
RG	Graves	Revision	Writer	Reflection	2	20	80
RHA	Applebee	Survey	9,11	Observation	309	130	81
RHGC	Grobe	Quality	8,11	Measurement	1750	11	81
RI	Stotsky	Read/write	EJHC	Summary	--	16	83
RJ	Hillocks	Research	EJHC	Meta-analysis	--	38	84
RK	McCulley	Cohesion	11	Measurement	493	14	85
RL	Swanson	Resistance	Facul	Case Study	2	29	86
RMC	Chesky	Audi/know	11	Measurement	80	10	87
RML	Lutkus	T-unit	4,8,12	Critique	--	19	87
RMS	Söter	WAC	EJHC	Summary	--	14	87
RN	Reed	Instrument	EJHC	Critique	--	8	88
RO	McCann	Argument	6,9,12	Measurement	95	14	89

also included in the NCTE composition research report series, were two exemplary experimental studies of a specific technique in teaching composition, known as sentence-combining. Another report selected from the NCTE composition research report series to be included in this study was Applebee's national survey, which was repeated several times later on to provide consistent information delineating the writing practice of school children in America.

The other fifteen reports were all centerpieces selected from the different research journals of different periods. Beach's report was an example of experimental study of the effect of edition on writing. Chesky and Hiebert's report was an example of correlation study of prior knowledge and audience and student's writing. Cooper's report was an example of combining objectives and measurements in a well-designed program. Donlan's report was an example of a survey study of teachers' attitude toward teaching writing in the content areas. Graves and Murray's report was a unique example of journals written by a mature writer and a writing researcher. Grobe's report and Steward's report were contributions by two Canadian researchers of English composition based on an American researcher's work in the 60's on the measurement of maturity in student writing. Hillock's report was on a new and powerful statistical approach known as the meta-analysis

applied to a host of composition research reports. Lutkus's report was a comparative study of two different measurements for syntactic maturity. McCann 's report was concerned with the student's ability to write argumentative prose with different prior knowledge. McCulley's report was on the correlation of writing quality and coherence and cohesion. Reed, Burton, and Vandett's report discussed the use of two measurement instruments developed by Hunt and Daly and Miller. Söter's report and Stotsky's report were two examples of summary reports aimed at presenting the results of previous research and identifying the direction for future research in a specific area. Swanson-Owens's report was a case study of the attitude of teachers toward the implementation of writing across the curriculum.

From the above description of the selection of reports for this study, it could be seen that the selected reports

1. used various research methodologies, including experiments, surveys, observations, case studies, synthesis, journals, correlation studies, measurements, and meta-analysis;
2. dealt with various aspects of the writing activity, including class size, editing, sentence-combining, prior knowledge, audience, modes of writing, teaching methodologies, teacher attitude, writing across the curriculum, syntactic maturity, writing quality, composing aloud, coherence, cohesion,

writing process, revision, writing apprehension, and reading/writing relationship;

3. focused on the secondary level;
4. involved different levels of writers, covering subjects ranging from elementary school through college, including both students and teachers and a mature writer and a researcher of writing; and
5. spread out over a period of thirty years, giving a sense of development in the research of composition.

The above description, when examined as a whole, presented a multi-level and multi-phased picture leading to the adoption of the hierarchical model, which could be presented in several phases:

1. the school phase:
 - institute level
 - grade level
 - curricular level
 - student level
2. the methodological phase
 - meta-analysis
 - experiment
 - correlation
 - case study
 - observation
 - survey
3. the social phase

geographical status
community status
socio-economic status
motivational status

4. the writing activity phase

quality
process
purpose
length
mode
time

5. the linguistic phase

originality
rhetoric
organization
mechanics

6. the communicative phase

official
school
personal
drafting

In order to illustrate the projected hierarchical model, a visual image can be presented later and, in this case, a linear regression mathematical model can be adopted to predict outcome of writing performance (Arnold, 1992; Bryk and Raudenbush, 1992; Willms and Raudenbush, 1989).

When the hierarchy theory is more carefully defined, other mathematical models can be designed to study this theory (Burstein, Miller, and Linn, 1979; Lee and Bryk, 1989; Lockheed and Longford, 1989; Raudenbush and Bryk, 1986; Raudenbush and Willms, 1991; and Willms and Raudenbush, 1989.)

Purpose and Design of the Study

This study was designed to answer the following research questions:

Firstly, what is the basic picture of research of teaching composition at the secondary level in the United States that could be introduced to the Chinese professionals to enhance their teaching and research of writing in English as a foreign language? Granted all the differences in language, schooling, age, experience and culture between the two populations, the assumption made here was that achievements made in one population may be beneficially transferred onto the other population with some adaptations.

Secondly, what are the feasible research methodologies adopted in this specified field that can profitably be adapted into a Chinese context for all the social and cultural differences? Methodologies in social studies manifest a number of significant differences among researchers in China and the United States. Some of them

are differences in form, but many of them are differences in fundamental concepts. A major purpose of this study was to identify those research methodologies that could be transferred into a Chinese context and also to describe the factors that hamper this adaptation.

Thirdly, what are the prevalent trends and issues in the research of teaching composition that could directly influence and promote research in the same field in China? Even if China may not currently be engaged in much research in this area, this study expected to bring the Chinese professionals into contact with the present status of research in which their American colleagues were working.

There were several ways to conduct this type of study, each with its own strengths and weaknesses. When the previously mentioned research questions were taken into consideration, the author of this study decided neither to adopt a historical approach, nor to adopt a case study approach, but rather a thematic approach to the analysis of selected research reports. A historical review was necessary only as the background setup for this study, and a single case study of a small number of individual research reports would have lost the overall perspective essential to the purpose of this study. The thematic approach used in this study consisted of classification, analysis, comparison and contrast, and evaluation within a thematic structure. Within each theme, several research reports could have been

combined to demonstrate that theme significantly in a systematic way. Through this approach, the author of this study expected that the requirements of the declared research questions could adequately be met.

Therefore, this study consists of six chapters. Chapter I discusses the rationale for the study, the selection of the reports, the purpose and design of the study, and the definitions used in this study. Chapter II describes the history of the research of teaching secondary composition in the United States, the methodology and foci of research, research in teaching English as a second language writing, and the suggestion of the hierarchy theory in studying this field. Chapter III analyzes the experimental studies included in this project, focusing on such themes as the theoretical foundation, the hypotheses defined, the designs, the sampling procedures, the experimental procedures, the instruments used for measurements, the statistical treatments used, the conclusions drawn from the studies, and the adaptability into a Chinese situation. Chapter IV analyzes the descriptive studies included in this study, focusing on such themes as the theoretical foundation, the concepts defined, the observation and measurement techniques, the use of case studies, as well as their adaptability into a Chinese situation. Chapter V focuses on the technique of meta-analysis, delineating the development of the summary

techniques of research literature and the use of meta-analysis in this area. Chapter VI discusses the introduction of the hierarchy theory into the area of composition research and presented a brief program for Chinese professionals in their effort to teach and research English composition.

In Appendix A, a list of classified bibliography is offered for further study and reference by Chinese professionals. The other appendices include some of the original documents quoted for this study.

Definitions Used in the Study

The following significant terms were used in this study:

Analytical study: An academic effort to obtain a deeper understanding of a certain research field by means of classification, analysis, comparison and contrast, and evaluation with a thematic approach.

Descriptive Research: An academic effort to determine and report the way the subjects and the phenomena are by collecting data to test hypotheses or answer questions concerning the current status of the subject of the study (Gay, 1987, pp. 10-11, 18).

Case study: An in-depth investigation of an individual, group, or institution to determine the factors, and

relationship among the factors, which have resulted in the current behavior or status of the subject of the study (adapted from Gay, 1987, p. 542).

Causal-comparative study: An academic effort to determine cause-effect relationships between variables without manipulating the "cause" variable (adapted from Gay, 1987, p. 13).

Correlational study: An academic effort to determine whether, and to what degree, a relationship exists between two or more quantifiable variables (Gay, 1987, p. 18).

Ethnographic study: An academic effort over an extended period to obtain a deeper understanding of a large number of variables in a certain research field by means of participant observation of a group of subjects in a naturalistic setting from cultural perspectives relevant to the subjects under study, described variously as ethnography, qualitative research, case study research, field research, or anthropological research (adapted from Gay, 1987; Louis M. Smith, 1979; and Goetz and LeCompte, 1984).

Experimental study: An academic effort to obtain a deeper understanding of a certain research field by means of artificially designed modules and at least one variable manipulated and other relevant variables controlled with a view to identifying correlational effect on one or more dependent variables (adapted from Gay, 1987, p. 544).

Hierarchical structure: An artificial theoretical framework adopted to organize and describe complex subareas of a certain research field (adapted from O'Neill, DeAngelis, Waide, and Allen, 1986, p. 55).

Observational study: An academic effort to obtain a deeper understanding of a certain research field by means of close observation for desired data in a group of subjects in certain research related situations (adapted from Gay, 1987, p. 547).

Qualitative study: An academic effort to obtain a deeper understanding of a certain research field by means of such designs as ethnography, case study analysis, historical or document analysis based on the belief that reality is ever changing (adapted from Reichardt and Cook, 1979).

Quantitative study: An academic effort to obtain a deeper understanding of a certain research field by means of such designs as experimentation, survey analysis, standardized observational research, and simulation models based on the belief that the reality is fixed (adapted from Reichardt and Cook, 1979).

Research report: Published article describing the systematic and relevant academic effort resulting in a deeper understanding of a certain research field.

Secondary: Seventh through twelfth grade in American school systems.

Thematic approach: Research method which organizes research

activities by focusing on meaningful topics related and relevant to the objectives of the research project.

Themes: Any of the topics related and relevant to the objectives of the research project.

In conclusion, the purpose of this study is to offer a comprehensive view of the research in teaching composition in the United States and design a set of guidelines based on this view so that these guidelines could be applied by Chinese professionals both to their teaching of English composition and to their research concerning the teaching of English composition. The specific area for this study was defined as secondary composition. Different research methodologies and their different aspects were discussed under different themes in this study. A collection of twenty reports were hand-picked by the researcher for analysis. To further facilitate follow-up studies by Chinese professionals, this researcher has provided a selected bibliography of works cited in these twenty reports which are listed in Appendix A. The definitions used in this study for analysis are in their broadest senses. It is the researcher's sincere hope that this beginning effort could be followed by studies in other aspects of teaching English composition in the United States for similar purposes.

CHAPTER II

THE HISTORY OF RESEARCH IN TEACHING COMPOSITION

Early Research in Teaching Composition

During the nineteenth century and earlier, very little research had been done in teaching composition at the secondary level in the United States. By 1850, English was still a minor subject in the secondary school curriculum. As a matter of fact, English, as it is, did not exist until the 1890's in the United States. English composition, when taught, usually consisted of monthly or bimonthly essays written on abstract or philosophical subjects with little or no formal instruction, and was often linked closely to declamation and oratory (Judy, 1).

Beginning in the 1920's, a number of research studies were conducted concerning the teaching of English in the United States in the nineteenth century, though not specifically on teaching composition as the major theme. Of these studies, the most prominent works were the following:

Stout's The Development of High School Curricula in the North Central States from 1860-1918 in 1921

Lyman's English Grammar in American Schools Before 1850 in 1922

Hays's College Entrance Requirements in English: Their Effects on the Schools in 1936

McBride's The Teaching of English in the Southern Ante-Bellum Academy in 1941

Kitzhaber's unpublished doctoral dissertation
"Rhetoric in American Colleges, 1850-1900" in
1953

Stephen Judy's unpublished doctoral dissertation
"The Teaching of English Composition in
American Secondary Schools, 1850-1893" in 1967

Berlin's Rhetoric and Reality: Writing
Instruction in American Colleges, 1900-1985 in
1987

To find out the status of research in this field during this specific period, we could also look into the journals published during the second half of the nineteenth century. The major journals that carried studies in this field were American Journal of Education and Journal of the Proceedings of the National Education Association. After 1880, more professional journals came into being, such as Education, Educational Review, School Review, and The Academy.

Of the 119 monthly and quarterly periodicals that had already appeared in the United States when the first issue of Barnard's American Journal of Education came off press in 1856, fourteen were related to education. Actually, while not all the articles carried in these periodicals were research reports, these periodicals were the chief places where researchers in education could get their reports published. Of these fourteen periodicals, one was for special education, one was at the college level, and twelve were published by the different states for their schools. These states included Connecticut, Illinois, Indiana, Massachusetts, Michigan, New Jersey, New York, Pennsylvania,

Rhode Island, Wisconsin and Ohio, the last of which had two periodicals instead of one. These periodicals for state schools had a volume of eight pages of quarto to regular 96 pages. When Barnard's journal came into being, it had a volume of about 130 pages in average and, above all, it was supposed to be a national journal. Therefore, American Journal of Education became the leading journal of all the education journals during that period.

Contributors to these publications formed a wide spectrum. Secondary school teachers, principals, superintendents and other personnel, as well as university professors and administrators were among the most frequent writers. The issues raised and discussed in their articles represented the general interest of the time and many of the arguments offered in these periodicals became clarified and accepted in the later implementation of school curricula.

When The American Journal of Education of the nineteenth century was surveyed, it was clear that study about the teaching of English at that time was heavily loaded with knowledge about the language. One very prominent article in the first volume of the journal was by John Hart (1855) entitled "The Anglo-Saxon Element in the English Language." The entry of "writing" was not even included in the index. The first time "writing" was entered into the index of the journal as a separate entry was in volume 7, number 19, in 1859 when the journal published the

translation of the Swiss educationist Heinrich Pestalozzi's work under the title of "Selection from Publications by Pestalozzi," in which Pestalozzi presented writing as a method of instruction (Pestalozzi, 1859). Meanwhile, the legacy of statistics was strongly manifesting itself throughout the journal. In each volume, charts and figures could be seen carrying statistics of various sources related to education, concerning general population, school population, handicapped population, school personnel income, federal and state funding for schools, numbers of journals available, and statistics in the comparative studies of American education with education of other countries, especially those in Europe.

This century saw the growth and gradual maturity in the research of writing in the light of the growth and maturity as manifested in the empirical research of the natural sciences. During the first half of the century, there was a strong tendency for the research in writing to emulate the empirical research in the physical sciences, as it was generally believed that all types of research should follow the established pattern of a relatively more successful pattern of research like that of the natural sciences. This obsession dictated the direction of the research in writing for a long time.

During this period, journals that carried research reports in this field not only increased in number, but also

showed a strong tendency to imitate the experimental designs in the research of the natural sciences. Along with this tendency, an increasing number of studies following this type were carried out and the designs of these studies became increasingly precise. Along with the increase of the knowledge about empirical research, conditions for these research studies were delineated with increasing accuracy. Eventually, researchers came to a point where they believed that they had detected some limitations in this approach.

Another source of publication to look into was the pool of dissertations and theses by university graduates. In 1938, University Microfilms decided to publish annually dissertation and thesis abstracts available in microfilm form titled Dissertation Abstracts International (DAI). The first five volumes of this journal were titled Microfilm Abstracts. In 1952, the title of this journal was changed to Dissertation Abstracts, which was changed to Dissertation Abstract International in 1970.

Between 1938 and 1950, the field of education had 448 dissertations entered in this journal, and only four were remotely related to the field of writing: two were concerned with spelling (Dumas, 1942 and Scrafford, 1943); one was concerned with shorthand training (Dempsey, 1949); and one was concerned with test-study versus study-test method (Foster, 1943).

In 1957, the Soviet Union, which the United States

regarded as the chief competitor at that time on the world political as well as military arena, successfully launched the first man-made satellite and greatly startled the United States. Governmental and academic institutions responded immediately. Funds and regulations were provided so that the United States could catch up and overtake the Soviet Union in the near future.

When American educators began to examine the educational systems of the two countries, they detected the deficiency extant in the American school system. A national anxiety was displayed in the question "Why Johnny can't read." Survey experts in different fields came up with the same kind of report, and researchers began to develop better curricula to prepare American youth for a more challenging future.

It is this scenario that influenced the researchers to examine, among other school-related areas, the teaching of writing and these studies gradually made themselves felt in this academic field. Governmental reports, documents, books, journal articles, theses and dissertations, as well as research reports funded by all types of foundations, companies and organizations were presented, going into almost all aspects of school education; and the teaching of writing was certainly no exception.

The Education Resource Information Center, known as ERIC, has kept up with information of the publication of

these different types of documents. During the ten years beginning from January of 1966 through December of 1975, a total of 6,747 publications came off the press on writing alone, including books, journal articles and various types of research reports, of which 178 research reports were dedicated to the instruction of writing at the secondary level. During the next seven years beginning from January of 1976 through December of 1982, there was a total of 10,961 publications on writing, of which 464 research reports were dedicated to instruction of writing at the secondary level. During the last ten years beginning from January of 1983 through December of 1992, there was a total of 23,153 publications on writing, of which 1,404 research reports were dedicated to the instruction of writing at the secondary level. Besides the interest shown in such aspects of teaching writing as methodology, motivation, and evaluation, the research reports also went into such fields as curriculum design, writing across the curriculum, writing apprehension, writing as a process, writing instruction for the learning disabled, and computer assisted writing instruction.

In this combined effort, Braddock, Lloyd-Jones, and Schoer's book Research in Written Composition (1963); Hillocks's book Research on Written Composition: New Directions for Teaching (1986); and Lauer and Asher's book Composition Research: Empirical Designs (1988) became the

three most outstanding works that tried to summarize what had been done in the field of research on teaching composition and to envisage the direction in which newer efforts were likely to be guided.

Braddock, Lloyd-Jones, and Schoer's work highlighted five "most soundly based" research studies in teaching composition and discovered that most research studies had serious limitations. This study was considered to be so significant that everyone attending the NCTE (National Council of Teachers of English) convention in San Francisco in 1963 was given a copy of Chapter 2, which detailed how exemplary research in composition should be conducted. They declared that the age of "alchemy" in the research of teaching composition should come to an end and a more organized effort should be made based on a "scientific method." (Braddock, Lloyd-Jones, and Schoer, 1963, p. 5)

Twenty-three years later, Hillocks (1986) applied, in his study, a powerful statistical process known as meta-analysis to an enormous amount of research reports that appeared after Braddock, Lloyd-Jones, and Schoer's study. Hillocks included about 2,000 (Hillocks, Jr., 1986, p. xviii) research reports in his selection, identified some approaches in the teaching of composition as having different effect sizes. Then, two years later, Lauer and Asher (1988) published their work describing the spectrum of various research methods used by researchers in the teaching

of composition and illustrated the eight types of empirical research designs frequently used by different researchers.

Methodology and Foci of Research

With the development of research in the field of composition instruction, research methodology itself was also developing. Researchers first borrowed methodology from empirical science and showed great enthusiasm for it. However, difficulties seemed to grow as researchers tried to single out the individual variables for control purposes as is frequently the case with empirical studies in the physical sciences. Some scientists identified as many as 100 variables involved in research in the humanities (Nunnally, 1975). All these variables supposedly exerted a composite influence on the subjects under study. It was almost impossible to have perfect control in determining the effect of any individual variable. A second difficulty was that all actual writings occurred under numerous variables at the same time.

In front of the unconquerable difficulties, researchers logically turned to the development of the research methodology in other humanities and found another useful method known as ethnography. Ethnography was a method widely used by sociologists and anthropologists in studying unknown or unfamiliar cultures. Researchers of writing

instruction borrowed this method by assuming that the writing environment was a totally unknown subculture and applied the ethnographic research approach to the process of writing instruction and made some surprising discoveries. Research specialists who contributed to this methodology included Patton (1980), Bogdan and Biklen (1982), Dobbert (1982), Bereiter and Scardamalia (1983), and Goetz and LeCompte (1984).

Researchers who used ethnography entered the writing setting presumably without any preconceived theoretical structure of procedures about conducting research in composition. The techniques they used in collecting data should be as much devoid of subjectivity as possible. They did observations and recordings of what happened in the writing setting, trying to make note of everything within their perception. They dissected the huge amount of collected data into little pieces and classified and reclassified the data in an effort to determine what kind of patterns existed in the data. After that, a second round of pointed observing, interviewing and recording took place, as well as a second round of analyzing data. Then there was a third round, with more focused observations, interviews and recordings. On and on, this procedure went until the researchers believed that the data gleaned were sufficient. Then discussions followed on the enormous quantity of ethnographic data. Patterns were extracted from the data.

New directions for research were suggested.

The appearance of this group of researchers changed the perspective within the field of research in composition instruction. Scholars began to view their research methods with more options, which resulted in a broader perspective and more research techniques available.

The focus of research also showed signs of change. In 1971, Janet Emig published her NCTE Research Report No. 13, under the title of The Composing Processes of Twelfth Graders. "She called for the abandonment of the product as the primary interest and for the encouragement of process writing teachers can initiate through imagination and sustain through empathy and support." (Banister, 1991)

After Emig, the process writing paradigm developed both among researchers as well as classroom teachers. Of researchers, Mischel (1974); Stallard (1974); Perl (1979); Pianko (1979); Bereiter, Scardamalia, Anderson, and Smart (1980); Bridwell (1980); Graves (1980); Hayes and Flower (1980); Bereiter and Scardamalia (1982); and Matsushashi (1982) became very active in this endeavor. Their subjects differed widely, from six-year-olds in Grave's study to college students in Pianko's and Perl's studies, covering students of remedial, average and advanced writing abilities. A number of studies also involved professional writers or technical writers.

Most research studies using this paradigm examined such

variables as "prewriting behavior, activity during pauses, rate of writing, and what writers do when they stop" (Hillocks, 1986, p. 1) as well as writing apprehension. These variables were distinctly measurable from a behavioristic approach. When the thinking process was involved, the researchers required their subjects to "think aloud." Though this skill was far from perfect, it began to reveal the mental processes involved in writing. By using the audiotape or the videotape, most researchers tried to record as faithfully as they could what was going on in the process of writing for further reference and analysis.

Research in Teaching Writing in the Field of English as a Second Language (ESL)

One important aspect in the research of teaching writing is in the ESL (English as a Second Language) field. In the last thirty years, ESL professional associations, degree programs in colleges and universities, academic journals and magazines, and books on ESL as an academic field have emerged to such an extent that it can be very well said that now ESL has been established as an important and inseparable part of the age-old profession of teaching English. In the United States, the TESOL Association, established in 1966, is a nation-wide organization of ESL professionals with more than 23,000 members and 76 regional

associations. This organization published a number of journals specifically oriented toward ESL instruction, the more important ones being TESOL Quarterly (1967) and TESOL Journal (1991). Many American universities now offer ESL degree programs, with a few offering the terminal degree.

As a matter of fact, along with the development of the English language in history, the teaching of English as a second language also developed. The growth of the English speaking population mandated that the teaching of English be carried on in its various forms among speakers of various other languages. Now more than 750,000,000 people use English. Of these people, more than half use English as an official language, a second language or a foreign language (McCrum, Cran, and MacNeil, 1986, 19). The ESL instructional task is enormous. When the present status of ESL instruction was examined, it left little doubt that research was needed to improve this global educational undertaking.

Silva (1988) claimed that studies in ESL composition started about 1945, the beginning of the modern era of second language teaching in the United States. In 1945, Fries published Teaching and Learning English as a Second Language, a book which marked the starting point of modern ESL studies. It is true that the development of ESL writing borrowed heavily from studies in the teaching of English composition to native speakers of English in paradigms,

methodologies, and concepts. However, ESL studies have their own unique aspects which deserve attention.

ESL covers a wide range of instructional activities. According to Judd (1978), a flood of terminology and acronyms was used in the teaching of English to speakers of other languages (TESOL). TESL meant the Teaching of English as a Second Language. TEFL meant the Teaching of English as a Foreign Language. TEIL meant the Teaching of English as an International Language. TEAL meant the Teaching of English as an Auxiliary Language. TEIAL meant the Teaching of English as an International-Auxiliary Language. TELWC meant the Teaching of English as a Language of Wider Communication. EAL meant English as an Additional Language. EFL meant English as a Foreign Language. ESP meant English for Special Purposes. EST meant English for Science and Technology.

In comparison, ESL is both similar and dissimilar to EAL, ELWC, or EFL. They all have English as a target language. The mother tongue of the learner is anything but English. However, the purposes of these language teaching programs are very different, involving a continuum of amount of effort, intensity of practice, and scope of use on the part of the learner.

All these language teaching activities share many of the issues involved in the study of ESL, and yet possess their own unique issues, such as issues arising from the

differences in culture, linguistic structure of the mother tongue, geographical locations, economy, political structure, objectives, and other factors. Furthermore, ESL for different age groups, for different levels, and for different purposes are designed to meet the challenges arising from today's multicultural society.

Silva pointed out that so far four most influential approaches existed in ESL writing instruction: controlled composition, current traditional rhetoric, the process approach, and English for academic purposes. He emphasized also that all these and other approaches underwent a merry-go-round routine: limited formulation, enthusiastic promotion, uncritical acceptance, premature rejection, and takeover by a shiny newer approach (1988). In order to get out of this endless routine, it was suggested that a coherent model of ESL writing instruction be considered.

Distinctive ESL theories focused on second language acquisition. Krashen (1981), in discussing the theoretical model of second language acquisition, summarized past literature and presented the five hypotheses in L2 acquisition: the acquisition versus learning hypothesis, the natural order hypothesis, the monitor hypothesis, the input hypothesis, and the affective filter hypothesis. He also summarized the factors that influenced second language acquisition as second language aptitude, role of the first language, routines and patterns, individual variation, and

age difference.

Carrell and Eisterhold (1983) discussed the schema theory (Bartlett, 1932; Rumelhart and Ortony, 1977; Rumelhart, 1980) in ESL reading instruction, a psycholinguistic theory used in the study of ESL, which Goodman described as a "psycholinguistic guessing game." (1967) Each of these popular theories in the ESL profession attempted to account for one aspect of language learning. It seems to the author of this study that an overriding theory should and can be devised to account for both the discrete theories so far discussed in the ESL literature and a deeper theoretical structure that serves to put all the discrete theories in one unified structure. This theoretical structure is the hierarchical structure to be discussed more fully later.

In the past thirty years or so, there have been dozens of books published on ESL writing, hundreds of articles had been published since 1966 on ESL writing, about three hundred of which were based on research. Of the 303 research related articles, 102 articles were on college ESL writing. Research in ESL writing fell into several identifiable categories. These categories included second language (L2) acquisition, error analysis, age group differences, influence of first language (L1) on second language (L2), comparison between L1 writing and L2 writing in different scenarios, cultural input into ESL writing,

effect of teacher/peer feedback, ESL writing process, attitude in ESL writing, effect of different types of writing tasks, and grading of ESL writing.

The earliest dissertation on ESL was completed by Lum (1972) and the earliest dissertation on ESL writing was completed by Davidson (1976). Between 1976 and June of 1993, 185 doctoral studies had been conducted in the various aspects of ESL writing. These studies covered a wide range of topics, including the ESL learner of various descriptions, the ESL writing instructor, different variables involved in ESL writing instruction, different aspects of the ESL learning process, and different ESL methodologies.

The research methodologies used in these studies were varied as well. A large number of studies were descriptive studies, attempting to describe the subjects of the studies, the writing behaviors and products of various ESL populations, including studies based on descriptive statistics and case studies based on intensive and prolonged observations. Frequently analysis was incorporated, including comparison and/or synthesis. Some studies were experiments conducted to determine correlations between roughly defined variables.

One of the major foci in these dissertations was on ESL learners of a specific mother tongue, like Spanish (Giffen-Glad, 1992; Jordan, 1992; Lopez, 1990; Mooney, 1991),

Chinese (Chen, 1992; Song, 1992; Teng, 1991; Wang, 1992), Korean (Hwang, 1992, Lee, 1991), and Malaysian (Morgan, 1990; Mustapha, 1992). Another major focus was on ESL learners of different age groups, like elementary students (Ezer, 1990; Kim, 1991; May, 1992), secondary students (Song, 1992); college students (Dessner, 1991; Hafernik, 1990; Kitchin, 1991; Smoke, 1991), graduate students (Xu, 1990), and adults (Martin, 1991).

Besides interest in ESL learners of different descriptions, attention was also paid to the instructional methodologies, including situational simulation (Li, 1992), whole language instruction (Chen, 1992), the integrated activity learning sequence approach (May, 1992), group interaction (Haber, 1991), metacognitive strategies (Kim, 1991; Sanchez-Villamil, 1991), and computer assisted writing (Dziombak, 1991).

While more fundamental research was conducted to determine the variables involved in ESL programs (Adipattaranun, 1992; Mooney, 1991), some studies were concerned with a certain aspect of the ESL learning process, like homework patterns (Giffen-Glad, 1992), measurement of giftedness (Jordan, 1992), revision (Lam, 1992), writing apprehension (Hayward, 1991), textbooks (Song, 1992), product (Ferris, 1992), and enrollment for ESL programs (Allameh, 1990).

The above profile of books and articles published and

the recent dissertation studies conducted shows that research interest in the field of ESL writing instruction has been developing fast, gradually going into the multitude of aspects explored in L1 writing research in English. This connection, on the one hand, revealed the relationship between L1 writing research and ESL writing research, and, on the other hand, suggests that an overriding hierarchical structure is justifiable for theoretical purposes.

Toward the Hierarchy Theory

With all the development of the research methodology in the field of composition instruction, the status quo of this academic endeavor, nevertheless, was still so much in need of organization and direction as well as theoretical structure that Gary Hatch called it a "crazy quilt." (Hatch, 1991) When the existing theoretical structures of composition research were examined, a number of theoretical structures could be identified, of which the theory of classical rhetoric, behaviorist theory, writing as thinking theory, and computer emulation theory were the major milestones along the path of its development.

Classical rhetoric put emphasis on the product of writing. Adherents to this theory (Bain, 1867; Connolly, 1953; Rockas, 1964; Bailey, 1965; Corbett, 1965; Young and Becker, 1965; Martin, Ohmann, and Wheatley, 1969) studied

the elements of the written composition in all its forms, advocated the use of model writing as prototype, and promoted ingenious application of elements singled out from rhetorical studies in the instruction of writing.

Behaviorism, borrowed from psychological studies, put emphasis on observable and measurable behaviors involved in the act of writing. Adherents to this theory studied the observable and measurable elements of the writer's behavior, advocated interaction between writer and human and/or non-human environment, and promoted motivated writing. The assumption of these researchers was best expressed in Skinner's statement "We are on safer and more promising ground if we stick to the behavior." (Skinner, 1978, p. 18)

The writing as thinking theory (Vygotsky, 1934/1962; Bruner, 1971; Luria and Yudovich, 1971; Martin, 1976; Emig, 1977, 1983; Applebee, 1981; Hays, Roth, Ramsey, and Foulke, 1983), also borrowed from psychological studies, put emphasis on active use of the mind known as the Black Box. Adherents to this theory studied many ways to stimulate the working of the mind in generating and arranging ideas, advocated critical thinking techniques, and promoted creative writing.

Computer emulation theory, borrowed from artificial intelligence studies in modern technology, put emphasis on the cerebral functioning of input, storage, processing and output of information in the writing process. Adherents to

this theory studied the cerebral schemata, advocated problem-solving techniques by exploring and exciting cerebral activities and critical thinking skills, and promoted organized writing.

These theoretical structures revealed the nature of writing with increasing insight and cumulative focus. Studies of what is put down on paper, studies of how writers put down writings in observable and measurable behavior, studies of how writers generate what is to be put down on paper, and studies of how the brain, like a computer processor, generates what is to be put down on paper--all these studies delineate the path along which knowledge was gained step by step about one of the greatest mysteries of human behavior. For the same purpose, the author of this dissertation suggested a hierarchical structure based on previous studies in writing as well as in other areas of study for further exploration of this field of learning.

The extremely complex nature of writing and the even more complex nature of instruction of writing are still defying the human effort in discovering the basic theoretical structure for both interpreting and completely controlling the act of writing per se. However, as Hillocks, Jr., put it, "it has come a long way," (Hillocks, Jr., 1986, p. 61) and this study was meant to be part of the collective effort in furthering the gigantic human enterprise to decipher this mystery ultimately.

With the increasing overlapping among different fields of studies, the study of writing has also extended into a large number of areas of study. Numerous facts have pointed to the organized structure of the universe as well as of the human activities in the universe so that a more global understanding has become necessary in understanding individual areas of research interest. True as it is that most branches of learning have been advanced through analytical approaches, the need is now overwhelmingly demanding that scientists also follow synthetical approaches in applying what has been learned in analytical studies and furthering inquiry in the same research area. As a matter of fact, the act of writing so heavily involves psychological processes, physiological processes as well as temporal and spacial factors that its communal and social levels are frequently ignored. As it is, the human act of writing can be studied at a variety of levels, with some levels explored more thoroughly than others. These levels seem to have some kind of structure. Also within each level there are a number of sublevels, similarly, with certain structures. To illustrate this point, the hierarchical model is designed such that this hierarchy is a complex structure which not only allows for different levels to be arranged in this model but also permits each level to embody different subhierarchies, and, additionally, which contains not only the existence of substance, but the

dimensions of time and space as well.

In the past five decades, many authors have pointed out the relevance of hierarchical theory to the study of complex systems (O'Neill, DeAngelis, Waide, and Allen, 1986, p. 55). As early as 1942, Egler used the notion of an arbitrary hierarchy in explaining the ecosystem. In 1945, Novikoff used the concept of hierarchy in the study of biology. In 1954, Feibleman used the hierarchy theory to examine the methods of explanation in science. In 1969, Whyte, Wilson, and Wilson discussed the long history of the hierarchy theory in the history of science. In 1970, Mesarovic, Macko, and Takahara formalized the hierarchy theory into a mathematical theory. In 1975, Nunnally suggested the use of the mathematical model for the hierarchical analysis in humanities. In the past half century, numerous scholars either studied the hierarchy theory itself or used the hierarchy theory in dealing with the extremely complex areas of study they were involved in so that today it is possible to use the developed hierarchy theory to study the extremely complex field of writing and the teaching of writing.

Pattie pointed out that we normally use the hierarchy theory to deal with complex systems (1973). The hierarchy theory, according to Mesarovic and Macko (1969), could resolve the dilemma, inherent in the study of complex systems, between simplifying details and accounting for observable behavior. Simon (1962) declared that the

hierarchy theory could decompose medium-numbered systems by focusing on an internal redundancy of structure. Later, he also argued that the hierarchical system was associated with a fundamental parsimony in the interaction among components (Simon, 1973). Therefore, it was made possible to isolate dynamics at a single level and ignore details at lower levels of this hierarchical system. In this system, according to Rowe (1961), we normally look to higher levels for significance and to lower levels for mechanisms, and, according to O'Neill et al, each level of the hierarchy corrects for errors or inefficiencies in the levels below (O'Neill, DeAngelis, Waide, and Allen, 1986, p. 58).

Another important factor related to the hierarchy theory is the measurement. Rosen (1977) tackled the issue of measurement in the hierarchy theory and pointed out that past knowledge was based on past observation done with past instrumentation, which did not include all possible measuring instruments, and, therefore, we could not consider any one description based on past measurements to be absolute or necessary. In other words, the degree of perfection in instrumentation has the most important impact on the study of complex systems.

Thirdly, the hierarchy theory has a multi-dimensional quality, involving different levels of systems as well as the spatiotemporal scales over which we observe complex systems of phenomena. Over-simplification of the hierarchy

theory should be avoided.

In conclusion, a brief review of the history of the research in English composition at the secondary level was offered. A number of important studies in this area were listed. Besides, this chapter discussed the several theoretical frameworks frequently adopted in the research of teaching composition and offered the hierarchy theory for further discussion.

CHAPTER III

EXPERIMENTAL RESEARCH IN TEACHING WRITING

Before Braddock, Lloyd-Jones, and Schoer's committee made its final decision about including five particular studies to analyze in its report (coded as RA), their team screened more than 1,000 bibliographic citations in 1961. In 1962, groups of experts were consulted and more items were included in this process. By the end of 1962, a final decision was reached that the following five studies, of which four were by American researchers and one by a British researcher, would be included in the final report for detailed summary by this committee: the Smith study (1931), the Kincaid study (1953), the Buxton study (1959), the Becker and Dallinger study (1960), and the Harris study (1962). Though not an experimental research report in itself, RA made it a point to include in its report for detailed summary only the "most soundly based" experimental studies.

For their selection, Braddock, Lloyd-Jones, and Schoer identified about thirty bibliographical sources for studies conducted from the 1920's up to the early 1960's (Braddock, Lloyd-Jones, and Schoer, 117-118), and listed a total of 504 items in their bibliography attached to the report. The researchers decided not to include in its final selection studies (1) that were not empirical research at all, what might best be termed "pilot" or "local" studies, (2) that

seemed to be incompletely described, and (3) that were based entirely on objective tests (p. 119).

From the criteria offered by RA, it could be seen that the definition of the "most soundly based" study was a definition of "what is not." To turn this "what is not" definition into a "what is" definition, the "most soundly based" study should be interpreted as a study that had a well defined null hypothesis; that had a true experimental design which had (1) a high level of validity and a high level of reliability, namely, with pre-tests and post-tests that measured what was meant to be tested in the samples of subjects and was consistent in results; (2) subjects who were randomly assigned to each group; and (3) a desired level of significance that was previously decided. These were the basic rules to be observed in a well-designed experimental study originally in natural sciences. The five studies selected in RA were all of them in agreement with these criteria.

A look at Table 1 also revealed that the O'Hare study (1971, coded as RBO), and the Beach study (1979, coded as RF) also belonged to this category. This chapter was to discuss the themes that occurred in the above mentioned seven experimental studies. For convenience of narration, the five studies were also coded: the Smith study was coded as RAS, the Kincaid study as RAK, the Buxton study as RABX, the Becker study as RABK and the Harris study as RAH,.

Philosophies, Theories and Hypotheses

The philosophy of the experimental study dated back to very ancient times. As early as the 4th century B.C., Aristotle (384-322 B. C.) already pointed out the order and regularity of the material world (1899). However, it was Francis Bacon (1561-1626) who was credited with developing the modern scientific movement, as he was the one who maintained the thought that stressed not only an understanding of the material world but control of it as well. Based on this thought, scientists began to talk about the "laws of nature." (Ozmon and Craver, 1981, p. 54)

Theories formulated through these philosophical considerations were very rigorous in nature. In the light of the materialistic philosophy, perception of the material world could be obtained through the human senses, and understanding and knowledge of the material world could be achieved through the function of the human mind. In addition, control could be exercised over the material world to facilitate changes in the material world so that these changes could be made to meet the needs of human beings at least for survival (p. 57). As the material world had its own orders, regulations, and laws, these orders, regulations, and laws could be observed and measured by instruments adopted by human beings for defined observations and measurements. Because of the existence of

these orders, regulations, and laws in the material world, the observations and measurements made by human beings would remain the same if the same conditions were applied. Therefore, the school of empiricism assumed that if exactly the same conditions were applied, the observations and measurements made on the subjects under study could repeat themselves.

Following this theory, empirical studies were developed in various areas of study, especially in the physical sciences, with brilliant successes. This proven rate of success, this professed rigor of implementation, and this declared efficacy for research in a combined manner made this approach so desirable that, in imitating and adapting, the researchers of areas of study other than the physical sciences sometimes lost sight of the differences between the characteristics of the mass and those of the mind in their effort to transfer what had been successful with the study of the mass onto the study of what was related to the mind. In the surge of this tide, the empirical research was believed to be all-important, all-capable and all-enlightening, whether it was applied to the study of the mass or to the study of the mind alike.

In its original application, empirical experiments were supposed to be designed to reveal, theoretically, what was not, or, in technical terms, the degree at which a null hypothesis was not to be rejected. Numerous efforts had

been dedicated to the study of this research approach. As many researchers contributed to the various forms of empirical research designs, many researchers studied the controls to be exercised on the empirical research designs, as well as the statistical procedures to be applied to the empirical research designs. This process was to be a long and tortuous one. Up to now, scientists are still working to define a perfect empirical research design. The more areas of study were involved, the more subtle the distinction between the mass and the mind became, and the more difficult the mission became to design a perfect empirical research format that would fit the needs of all research areas.

The five studies selected in RA (RABX, RAH, RAK, RAS, and RABK) and the other two studies selected for this study (RBO and RF) were all designed in the light of this school of philosophy. The criterion "most soundly based" which Braddock, Lloyd-Jones, and Schoer used in the selection of studies for summary was also in line with this school of philosophy, the major apparent discrepancy being the fact that the foci of these studies were strongly concerned with the function of the mind, and, in addition, the control of the studies were far from ideal.

At the beginning of RA, Braddock, Lloyd-Jones, and Schoer quoted a colleague who wrote, "What is the sense of attempting an elaborate empirical study if there is no

chance of controlling the major elements in it?" (Braddock, Lloyd-Jones, and Schoer, 1963, p. 5) This was a very powerful argument, which, regretfully, was not addressed further on, leaving critical readers wondering why this essential concept was thus ignored. By using the term "the most soundly based," RA referred to the issue of whether it was appropriate to disregard the lack of control of major elements in experimental research designs not as a dichotomy, but as a continuum, in order for the validity of such designs to be accepted. The studies carried out along this line attempted to imitate the successful pattern of research in physical sciences, taking the workings of the mind to be similar to the workings of nature, and assuming that no difference or insignificant difference existed between the two.

RA, in its effort to observe the criterion for selection, also identified a number of variables: the writer variable, the assignment variable, the rater variable, and the colleague variable, for the purpose of control in studies. Later studies added to this list numerous variables that needed to be controlled in experimental studies on writing or on the teaching of writing. When researchers found that it was impossible to achieve laboratory rigidity in control, randomization became the best single available and acceptable tool for statistical purposes (Gay, 1987, p. 276), though its ultimate legitimacy

in experiments was as yet still quite uncertain. First of all, randomization took on various forms for various situations, leaving a large area for debate as to whether the conditions defined for the application of a certain type of randomization were necessary and sufficient. Secondly, there still might be some critical variables unidentified in using some types of randomization.

When all of the above issues were scrutinized, a large gray area presented itself, where the quantified results achieved in these studies were, after all, more or less ambiguous. Furthermore, when the hypotheses formulated in these studies were examined, another aspect added to this gray area, making the whole situation even fuzzier.

In RABX, for example, the defined purposes of the experiment were to examine the differences between two instructional methods for composition, respectively termed as "freedom from restraint" and "prevision, writing, revision," which were called "two divergent views toward the teaching of composition." (Braddock, Lloyd-Jones, and Schoer, 1963, p. 58-59) The researcher attempted to find out the effect of one teaching methodology characterized as "freedom from restraint" and one teaching methodology characterized as "prevision, writing, revision." A serious conventional concern that frequently arose among researchers was whether the treatment applied to the experimental group was really what the discussion said it was. In the case of

RABX, "freedom from restraint" was one of such concerns. If the definition of "freedom from restraint" was to be necessary and sufficient, the treatment applied to the experimental group was only one type of "freedom from restraint," or only a few representative samples of "freedom from restraint" at the best. There were other types of "freedom from restraint" totally ignored in this experiment, like freedom from restraint on time, freedom from restraint on place, freedom from restraint on deadline, freedom from restraint on requirement, and freedom from restraint on criteria, just to name a few. If this "freedom from restraint" was true, any kind of restraint imposed on this group of subjects would invalidate the experiment.

Likewise, similar questions could be asked of the terms "prevision, writing, and revision." When the experiments were closely examined, it seemed that the "prevision, writing, revision" included "encouragement," "outlining," "choosing words and illustrations carefully," "warning against using unqualified and fallacious statements," "developing good unity within paragraphs and transition between them," besides a different style of marking and reviewing the paper (p. 61). At least, in this case, the terms used had a fuller content base. This group of subjects had much more classroom attention than the previous group of "freedom from restraint." The reader might be wondering what was taught to the subjects in the group

"freedom from restraint" during regular classroom instruction time? Or was the difference detected really the difference between "teaching" and "non-teaching" rather than the difference between "freedom from restraint" and "prevision, writing, and revision?" The point was whether the activities designed for the groups of subjects really showed the difference effected between the two methods thus described since the report declared that one was a "method in which no interlinear or marginal marks were employed on the student papers, no grades were noted on the papers, the only remark was a paragraph of generous comment at the end of the paper, and the students were not advised to correct or revise their papers," and the other was a method "in which the papers were thoroughly marked, graded, commented upon (adversely, when necessary) in a paragraph at the end of the paper, and discussed by the raters and revised by the students during 35 to 50 minutes of class time when each paper was returned," (p. 58) The discrepancy between the appropriate level of generalization of the actual activities designed for the experiment and the generalized term used in discussion and conclusion for those same activities in the experiments could be misleading.

The concept "two divergent views toward the teaching of composition," on the other hand, posed another type of bewilderment. What were the two views? How were the two views defined? Were the implications of the two views

totally antagonistic in nature or were they supplementary to each other, or a combination of both? Furthermore, were the two views accurately represented at the right level of abstraction by the two types of activities designed for the two groups of subjects put to test? With the results of the experiments, what conclusions were the appropriate within the known scope of human knowledge and what conclusions were made one or two steps too many away from the appropriate position?

Apparently, it was a very common case where terms of different levels of generalization were used interchangeably in experiments and in discussion. This situation posed an opportunity where the hierarchy structure had to intervene. When the term at a higher level of generalization was used, were the conclusions drawn from a lower level sample valid for this higher level term? If we were doing an experiment on fish, would the conclusions drawn on fish be equally valid on the level of animal? Empiricism warned the learned world about the process of inference at different levels of generalization. This point would be equally valuable for generalizations in educational empirical studies. Extra caution should be exercised on the validity of generalizations for terms of different levels of abstraction.

When the theoretical constructs of these experimental studies were examined, it became clear that different

studies were developed around different learning theories. For example, the O'Hare study used the learning theory of "mental discipline." When talking about the benefit of repeated practice on possible future application, O'Hare used the image of football players practicing "hundreds of plays many times so that at the right time, in the right situation, a dozen or so of these moves will have become both appropriate and habitual." (O'Hare, 1971, p. 31) The theory underlying this image was evident: an exercising mental muscle would perform the trained moves under actual circumstances in the future. The Beach study examined the difference in effect of between-draft teacher evaluation versus student self-evaluation. Beach assumed that the intervention of a carefully planned teacher-evaluation would significantly change the overt behavior of a student's revising ability. Apparently, theory of human growth and development was involved in this study. When the discussion of the Beach study was examined, what Beach found to be unaccounted for might be accounted for in the light of Bloom's taxonomy of educational objectives. The concern in RF was about why, of the four quality ratings, only support and degree of change showed significant difference, but not sequence or focus. Besides Beach's contemplation about teacher's appropriation of instructional attention and time, another factor might be students' ability to follow the instructions. "Support" was easier to comprehend and apply

because it involved addition of details which might be related to the topic in interest where the student had to decide only whether the details did support the topic, while "sequence" and "focus" might be more difficult to comprehend and apply because they both involved elements that needed to be evaluated against some criteria. For example, the students had to know in what sense some of the sequencing or focusing techniques were to be used, by what kind of organizational skills these details could be added, how many options they had, and which was the best choice for their specific writing assignment. These considerations involved a much higher level of thinking skills. It was not clear from the report what criteria were actually developed for the students in teacher evaluation process and what criteria were developed for the raters in determining the score so far as "sequence" and "focus" were concerned. Nevertheless, studies in the cognitive domain, which displayed the hierarchy structure, could be incorporated into the Beach study for further discussion.

The process of hypothesizing in experimental studies was another aspect of interest in this study. In RAS, a study exploring the effect of class size upon the efficiency of instruction in ninth-grade English in 1931 did not hypothesize. In the process of this study, the teacher-investigator gave a series of well-designed testing instruments of that time, including the following:

1. the Army Alpha,
2. the Haggerty Intelligence Examination, Delta II,
3. the Pressey Senior Classification Test,
4. the Terman Group Test of Mental Ability, Form A,
5. the Haggerty Reading Examination, Sigma III, Form B,
6. the University High School Grammar Test,
7. the ninth-grade minimum essentials spelling list for the state of Minnesota,
8. a composition rated according to the instructions in the Hudelson Typical Composition Ability Scale,
9. Briggs English form,
10. Kirby Grammar Examination.

Besides, a number of classroom tests and some questionnaires designed by the teacher-investigator were given. Results of these carefully designed tests were compared between a class of 51 students and a class of 21 students in the University High School at the University of Minnesota 1925-1927. This study was an early example of educators in English classes making efforts to investigate some puzzling issues by borrowing from the experimental research designs of the physical sciences so familiar to us now. The data reported in RAS were raw data from the test

scores. No statistical treatment was visible.

The RAK of 1953 studied whether a single paper written by a student on a given topic at a particular time constituted a valid basis for evaluation at Michigan State College. Hypothesizing in this study was obvious. The first hypotheses in this study were as follows:

1. Any given topic provides the same stimulus as any other topic.
2. Any given topic elicits constant responses at different times.
3. The psychological pressure of the examination situation has no adverse effect on the quality of student writing.
4. The quality of student writing is stable from topic to topic and from time to time, with or without the pressure of an examination, regardless of variations in writing ability."

(Braddock, Lloyd-Jones, and Schoer, 1963, pp. 84-85)

Each of the first three hypotheses focused on one independent variable only, namely, the variable of topic for the first hypothesis, the variable of time for the second hypothesis, and the variable of existence of examination pressure for the third hypothesis. For the fourth hypothesis, the independent variable was a collection of three: the topic, the time, and the pressure of an examination (p. 84). All the hypotheses formulated in this

study observed the rigorous rules defined in empirical research of the physical sciences in terms of null hypothesis formulated, single independent variable identified, and direction of dependent variables suggested.

The RABX study of 1959, also discussed earlier, did not use the null hypothesis approach. Instead, RABX stated its research purpose as to:

determine whether or not "regular practice in writing over a period of seven months would result in a significant improvement in writing skill; and which of the following methods was superior in improving writing skill: (a) a "freedom from restraint," or "Writing" method in which no interlinear or marginal marks were employed on the student papers, no grades were noted on the papers, the only remark was a paragraph of generous comment at the end of the paper, and the students were not advised to correct or revise their papers, and (b) a "prevision, writing, and revision," or "Revision," method, in which the papers were thoroughly marked, graded, commented upon (adversely, when necessary _ in a paragraph at the end of the paper, and discussed by the raters and revised by the students during 35 to 50 minutes of class time when each paper was returned.

(Braddock, Lloyd-Jones, and schoer, 1963, p. 58)

In contrast, RABK used the null-hypothesis approach. RA referred to the null hypothesis approach in RABK only in the part entitled Result, where it was reported that the investigators decided before the experiment "to use the 5 percent level of confidence as the level at which to reject the hypothesis of no difference.

The way of using a research question instead of a null hypothesis in formulating an experimental research design

runs the risk of accepting, when the designed level of significance is met, the opposite of the implied null hypothesis, which might not be true, either. There might be some other unidentified variables at work than the named variable under test. When this type of empirical study was carried out in the areas other than physical sciences, this risk was even likelier to occur.

The construction of hypotheses should, therefore, be very carefully done. In forming hypotheses for experimental studies, it would be desirable to go to lower levels of abstraction first, test out as many phases as possible at that level, and, based on knowledge obtained at the lower levels, gradually move upward in the hierarchy. It would also be desirable to use the null hypothesis approach in designing this kind of true experiment.

Designs, Samples and Procedures

When the theoretical structure of a research project had been determined, the design of the experiment could be of vital importance to the validity of the study. In RAS of 1931, the purpose of the research was to determine the effect of class size on the efficiency of instruction in ninth-grade English. The researcher used only two classes taught by the researcher herself. The investigation was conducted for two successive years. The technique of

matching was used in this experiment. Twenty-one students were carefully selected from the large class of 51 for the year of 1926-27 to match the 21 students of the small class, based on the results of a battery of intelligence tests, tests in reading, grammar, and spelling, as well as on the chronological age of the students, and a written composition. The other thirty students were used only to build up the size of the class. Some unique techniques were developed by the researcher in the instruction of the large class during the experiment, such as group activities, change of seats, frequent tests and exercises, peer grading and editing, class tabulation and analysis of errors in tests, use of mini-session activities, and stimulation of self-competition by asking the students in the large class to record their scores regularly.

The results of this experiment strongly indicated that "the teaching of composition to superior ninth graders (with an average IQ score of 116) may be generally as effective in large classes as in small classes if the teacher has the requisite ability and is given adequate time for preparing lessons and adequate assistance in handling the routine aspects of class management." (Braddock, Lloyd-Jones, and Schoer, 1963, p. 106) Looking back at the total design of the experiment, a careful reader might ask this question: Were the final results of the experiment an effect of the difference in the size of the class, or an effect of one or

a combination of the unique techniques introduced into a large class, or an effect of both?

Theoretically, an experiment to decide the effect of the size of class on the efficiency of instruction would have the other variables controlled. When the Smith study was examined, the variables at work in this experiment included not only the difference of the size of class, but also the difference of instructional skills at the same time. If the purpose of the study had been to determine the effect of a unique combination of instructional techniques on a large English class of 51 superior ninth-grade students as compared with a small English class of 21 students with a different approach to instruction, who were matched in terms of intelligence, chronological age, language proficiency, and written ability, the design of this study would have been better suited. However, if the purpose of this study was redefined as suggested, it could be seen that the effect of class size was buried somewhere in this experiment. How could the effect of class size be shown to have affected the results of the students in this case?

On the other hand, the Kincaid study of 1953 used the null hypothesis approach and adopted a design that was fairly close to the designs widely accepted in the research of physical sciences (Table 2).

So far as the design was concerned, this design was based on dichotomous variables under examination: similar

Key: Topic 1--"Distance Between Classes"
Topic 2--"Cost of Textbooks"
Topic 3--"Giving Directions"
W/O--Writing without the examination pressure
W --Writing with the examination pressure

	First Day Similar Topics		Second Day Similar Topics	
	Topic 1	Topic 2	Topic 1	Topic 2
Group A	W/O	W/O	W/O	W/O
Group C	W/O	W/O	W	W

	Dissimilar Topics		Dissimilar Topics	
	Topic 1	Topic 3	Topic 1	Topic 3
Group B	W/O	W/O	W/O	W/O
Group D	W/O	W/O	W	W

topics versus dissimilar topics, writing with the examination pressure versus writing without the examination pressure, and the first day versus the second day for writing. With the help of this design, the experiment was conducted in such a way that different aspects of these variables could be singled out for statistical treatment. Under ideal control condition, this design would be able to provide more focused and individualized data by its format. This format was most appropriate for testing a small number of basic dichotomous variables at the same time. If these variables were closely examined, it could be seen that these variables were at lower levels of complication. The time

variable focused only on whether it was the first day or the second day for the writing assignments. The topic variable focused only on whether the two topics for writing assignments were of a similar or a dissimilar discourse type. The pressure variable focused only on whether there was any examination pressure on the writing assignments.

The O'Hare study adopted the pre-test-post-test control group design described by Campbell and Stanley (Campbell and Stanley, 1971, p. 13). The design took the following form: (Braddock, Lloyd-Jones, and Schoer, 1963, p. 36)

R	O	X	O
---	---	---	---

R	O		O
---	---	--	---

Many studies included in this analysis had as their independent variables factors far more complex than the variables cited in RAK. RABX of 1959 dealt with the "method" variables. One was the method of "regular practice in writing over a period of seven months." One was the method "in which no interlinear or marginal marks were employed on the student papers, no grades were noted on the papers, the only remark was a paragraph of generous comment at the end of the paper, and the students were not advised to correct or revise their papers." And one was the method "in which the papers were thoroughly marked, graded, commented upon (adversely, when necessary) in a paragraph at the end of the paper, and discussed by the raters and revised by the students during 35 to 50 minutes of class

time when each paper was returned." (p. 58) RABK of 1960 also dealt with the "method" variables: one method of "making one instructor responsible for the training of his students and meeting them in class four times each week," one method of "making the student more self-reliant and independent by reducing from four to three the number of formal class meetings per week, by presenting the basic principles of the course ... in a bibliography of assigned and collateral readings, and by devoting most of the class hours to performances under the guidance and criticism of a classroom instructor," and one method of "presenting the basic principles of the course by 'experts' through the use of film recordings of television programs, supplemented by discussion of these principles and their application in performances under the guidance and criticism of the regular classroom instructor." (p. 107) RAH of 1962 likewise dealt with "method" variables. One method was known as "formal grammar" and the other method was known as the "direct" method. Similarly, RBO of 1971 dealt with the independent variable of "teaching a regular curriculum in English versus sentence-combining practice and a shortened version of this regular curriculum." (O'Hare, 1971, p. 36) RF of 1979 dealt with a more or less limited variable of "teacher evaluation versus guided student self-evaluation versus no evaluation between drafts." (Beach, 1979, p. 113)

When the level of complication of the variables under study was raised, the design of the studies tended to become less focused and exhibit an artificial enforcement of definition, resulting in a fuzzy process of conceptualization and a frequent deviation from the perspective originally possessed by the researcher. In such a case, the discussion of the experiment tended to be over generalizing. This phenomenon was typical in the hierarchy theory, as the higher the level of variable, the fuzzier the generalization concerning the variable would be, because when the variables covered more phases and dimensions, the generalization could be even more pervasive.

The size of the sample, the sampling procedure, and the randomization of the sample were other points that needed to be examined. The Smith study (1931) used a sample of 72 ninth graders in two existing classes, of whom 21 belonged to the small class, another 21 were selected in the large class to match the students of the small class, and the remaining thirty others were used to build up the size of the large class, in the 1926-27 school year at the University High School of the University of Minnesota (Braddock, Lloyd-Jones, and Schoer, 1963, p. 95). The Kincaid study (1953) examined "320 papers written between the fall term, 1951, and the winter term, 1952," by a sample of 80 "students (who were assigned to two classes based on the alphabetical order of their names during registration on a

daily basis, and further divided into four groups by counting every other student in each class) enrolled in the first quarter of the Written and Spoken English course (later named Communication Skills) at Michigan State College (now Michigan State University)." (p. 83) The Buxton study (1959) used a sample of 257 first- year students randomly assigned to six classes by using a table of randomized numbers at the University of Alberta during the academic year of 1956-1957 (p. 59). The Becker study (1960) used, as they finally appeared qualified for this study, 758 students, with 585 attending the two-semester course and 176 attending the one-semester course, who were assigned to three classes by the randomized hour of registration (p. 108). However, the report did not explain why 758 students were reported to have been enrolled in the experiment, whereas the students attending the two types of courses totaled 761. The Harris study (1962) used a sample of 228 students aged twelve to fourteen already divided, fortunately not according to ability, into two forms (classes) at each of the selected five London schools (p. 76). The O'Hare study (1971) used a sample of 83 seventh graders at Florida State University High School (O'Hare, 1971, p. 37). The Beach study (1979) had a sample of "103 students in three 10th grade and two 11th/12th grade classes at a high school located in a predominantly working-class neighborhood in Minneapolis." (Beach, 1979, p. 113)

From the above numbers, it could be discerned that the subjects in the experimental studies were all of them grouped in classes, of a close age range, of both sexes, of assumedly similar socioeconomic background, and of a size between 72 in the Smith study to 758 in the Becker study. Throughout the experiments, there were sometimes a few students coming into the sample or leaving the sample, who, as a rule, were not entered into the summary or the statistical treatment. The rate of absence of the subjects from class was not reported, however, making it difficult to assess what effect the rate of absence of the subjects would have on the final results, as most of the significant difference observed was attributed to classroom treatment.

The procedures of the experimental studies offered a variety of techniques for analysis. A survey of the procedures adopted in the seven empirical studies listed in this project revealed a number of interesting facts.

Most of the writing tasks were on a given topic in a predetermined school period (Braddock, Lloyd-Jones, and Schoer, 1963, p. 73). Smith did not disclose the time period actually spent by the average student in completing each of the assigned compositions for rewriting. Kincaid reported all four papers written during students' regular laboratory periods when they were accustomed to write (p. 85). Buxton reported a 50-minute period for original essay writing (p. 62). Becker reported a 450-word minimum, two-

hour expository theme besides a 45-minute objective test in English composition both for the pre-test and post-test (p. 111). Harris reported two 40-minute class periods for the writing of each of the two compositions (p. 73). O'Hare reported all the compositions written in class under teacher supervision during the first two weeks in October and the last two weeks in May without telling how long the students were allowed to complete their writing (O'Hare, 1971, pp. 46-47). Beach reported a fifty-minute period on Friday to write the rough draft and Monday and Tuesday class periods were spent on completing the final drafts (Beach, 1979, p. 114).

The topics for the writing assignments reported in the selected studies (RA, RBO, RF) included:

- RAS "Slatter's Hill Snowball Fight" (p. 98)
"The Most Exciting Ride I Ever Had" (p. 102)
- RAK "It Is (Is Not) Too Far Between Classrooms for Students to Travel During the Ten-Minute Period Allowed"
"Textbooks for Freshmen at Michigan State College Are (Are Not) Too Expensive"
"Give Directions to a Stranger at the Union Building Enabling Him to Get to Building A-6 on South Campus" (p. 84)
- RABX "What Music Means to Me"
"My Observation of Science Teaching in the Elementary School"
"My Visit to the Alberta Legislature" (and my conclusions)
"My Opinion of the Film 'Near Home'"
"Handwriting Can Be Improved" (p. 62)
"The Cold War"
"Profit from Friendship"
"My High School"
"The Significance of an Incident in History"

"Special Knowledge I Possess"
"Family Relationships" (pp. 111-112)

RAH "A Day in the Country"
"A Day at the Seaside" (p. 73)
"A House on Fire"
"A Foggy Night" (p. 75)

RBO A-1 Stranded in a Ghost Town
The Thing That Wouldn't Die
No Ordinary Forest
The Strangest Day Ever
B-1 Creature from the Lake
The Old Woman in the Fog
What an Unusual Day
Lost on Evil Island

A-2 The Dinner Table at Thanksgiving
A Rainy Night at the Football Game
Watching a Building Burning Down
The Amusement Park at Night

B-2 An Afternoon at a Fair in Autumn
A Cook-Out at the Shore
After a Bad Storm
Halloween Night

A-3 A lady in your neighborhood is not sure if she
should have you babysit. Convince her that
you can do the job.
Persuade your parents to raise your allowance
by a certain amount.
Convince your parents that you should be
allowed to go to the high school dance.
B-3 Imagine your parents have won a competition,
and the prize is a two week vacation, all
expenses paid. Persuade them to go to the
place of your choice.
Persuade a teacher to limit homework to one
night a week.
Convince your parents that you should be
allowed to decide how you dress.

A-4 In Dave Crockett's time transportation was
a very important part of people's lives.
People in those days walked a great deal, but
they also made use of horses, mules, coaches,
canal boats and ships to carry men and goods
from place to place. Imagine that a boy (or
girl) your age living in those days
accidentally entered the fourth dimension and

landed in Tallahassee this week. Naturally he would be fascinated by the changes in the means of transportation that have taken place since the frontier days. Write a report telling him of the many new means of transportation that have been invented since his day. Tell him how they work, what they can do and so on. Try to answer any questions you think he might ask.

- B-4 In Jim Bowie's time most frontier people's homes were log cabins, lit by oil lamps or candles. They had no running water and the few kitchen utensils they had were usually crude and simple. Imagine that a boy (or girl) your age living in those days accidentally entered the fourth dimension and landed in Tallahassee this week. Naturally he would be fascinated by the many changes that have taken place in homes since the frontier days. Write a report, telling him of the developments in the home since his day. Mention some of the many home appliances and gadgets that have been invented since Jim Bowie's time. Tell him how they work, what they can do, and so on. Try to answer any questions you think he might ask.

- A-5 Saturday afternoon at the movies
An ice-cream parlor
Sunday morning at church
A favorite quiet place
A backyard swimming pool

- B-5 A doctor's or dentist's office
The school principal's office
A new school on the first day
A traffic jam on a hot afternoon
Your room on a cold winter morning, when the heating system has broken down.
(one topic to be chosen from each group)

(O'Hare, 1971, pp. 102-105)

- RF 1. "'Most people are out for themselves.' Do you agree or disagree with this statement? Give reasons why you agree or disagree with this statement."
2. "Write about a well-known person, television/movie star, or friend that you like or that you dislike. Give reasons why you like or dislike that person."

3. "Write about one of your strengths and one of your weaknesses. Give reasons why you chose these as your one strength and one weakness."

(Beach, 1979, p. 114)

The time span of each experimental study usually depended on the level of complication of the hypotheses. In case of a lower level of complication of the hypothesis, the study could be carried out within one to several weeks, like the Kincaid study and the Beach study. If the hypothesis involved a full fledged "method" of instruction, the experimental study would normally last for one to two semesters or up to two years or even longer, like the Smith study.

While some of the studies indicated the qualifications of the most important factor in the treatment, the teacher, other studies like the one Beach conducted failed to give any indication of this factor, which could include such issues as who had been employed as the agent to conduct the treatment, what were the qualifications of the agent, what the motivation was for the agent in this treatment, what the agent's attitude was toward this experiment, whether the agent was carefully selected or randomly selected, what the criteria for selection of agent were, and how important the issue of the agent was in this type of experiments. The Smith study had as its agent the investigator herself. The Kincaid study only mentioned the fact that the papers evaluated in this investigation came from the classes of two

instructors (Braddock, Lloyd-Jones, and Schoer, 1963, p. 84). The Buxton study arranged "as nearly as possible, that no two classes in the same group were taught by the same instructor unless the instructor taught more than three classes," and that four instructors taught the English section of the three groups together with more than eighteen other instructors for the instruction of the total 11 courses in this study (pp. 59-60). The Becker study reported the agent as "a staff largely composed of part-time graduate assistants," who were randomly assigned with a number of restrictions to classes in this University of Iowa freshman program (p. 107). The Harris study reported that one teacher taught both forms of students in four schools and in the fifth school two teachers taught the two forms of students, and that "no certainty exists that each teacher was absolutely impartial in his attitude to the teaching or abandonment of formal grammar, or indeed equally adept at the one approach as at the other," but the investigator judged that "all the teachers were willing to be convinced by evidence, desirous of finding out the truth about the matter in hand, competent and practiced in the present grammar of the language." (p. 78) The O'Hare study specifically addressed the issue of the teacher, where two tables were used to demonstrate that the difference attributed to the factor of the two instructors was insignificant (O'Hare, 1971, pp. 58-59). The Beach study

did not mention this factor.

Compared with the agent variable, the rater variable received a much better consideration in the selected studies. The Smith study used five "experienced judges." (Braddock, Lloyd-Jones, and Schoer, 1963, p. 102) The Kincaid study selected nine out of more than forty staff members of the Department of Written and Spoken English at Michigan State College who "demonstrated a high degree of sensitiveness and reliability." (p. 86) The Buxton study used "a group of qualified readers appointed by the investigator," namely, graduate assistants specializing in the various fields with some "teaching experience," and "paid on an hourly basis out of funds allotted to this investigation by the Alberta Advisory Committee on Educational Research (p. 61). The Becker study used one "experienced" rater from the Communication Skills staff for approach, thus totally avoiding the "solely subjective judgement" in rating the compositions (p. 73). The O'Hare study offered the information concerning the eight evaluators used in the study, all of them "experienced English teachers attending Florida State University" at the time of the experiment, with an average of six years of teaching experience (O'Hare, 1971, p. 50). The Beach study used three judges who had "received approximately thirty hours of training in the use of the degree-of-change scale and the quality rating scales devised by the investigator."

(Beach, 1979, p. 114)

Extraneous variables were hardly controlled in the selected studies. Little effort was noticed in the selected reports even to identify this issue. One effort was made in the Kincaid study to prevent the intrusion of extraneous influences into the experiment by instituting a number of controls to allow for fatigue, motivation and the like. Aside from that the best effort made to minimize the effect of extraneous variables was to use some sort of randomization. Administrative limitations in schools sometimes mandated ready made classes, jeopardizing the validity of the generalizations made in the studies.

The treatment of the experimental groups in these studies was mostly individualized because of the individuality of the instructional processes. Hardly was it possible to reproduce the experiment without bringing in any noticeable changes in process, material, attitude, commitment, motivation, approaches, and any other variables that could be named. Moreover, the subjects could hardly be used more than once for the same treatment. Therefore, the possibility to replicate the experimental studies was very questionable in research studies of this type.

Instruments, Statistics, and Conclusions

What the researchers in these studies actually observed were numerical data. Different types of numerical data were extracted by means of different types of measuring instruments. How these instruments were designed, what these instruments were supposed to measure, how accurate these instruments and measurements were, how valid and reliable these instruments and measurements were, these were some of the major concerns of the researchers.

When the measured data were presented in the numerical form, a well designed branch of the mathematical science, namely, statistics, could be applied to the data thus gleaned and logical conclusions could be drawn by manipulating the numerical data through statistical procedures. Because of the experimental nature of the studies, the data were presented in both the pre-test and post-test score forms, so that the difference detected between the two sets of scores were supposedly attributed to the effect of the treatment.

When the actual procedures of the studies were examined, there were a few points that should be discussed in detail concerning the instruments. The Smith study used a number of well-designed tests as pre-tests for the purpose of matching the students in the small class with the same number of students selected from the large class. The IQ

tests used in the Smith study included the Army Alpha, the Haggerty Intelligence Examination, Delta II, the Pressey Senior Classification Test, and the Terman Group Test of Mental Ability, Form A. The reading ability of the students was tested by using the Haggerty Reading Examination, Sigma III, Form B. The students' grammar was tested by using the University High School Grammar Test. Their spelling was tested by using "a pre-test in the ninth grade minimum essentials spelling list for the state of Minnesota."

(Braddock, Lloyd-Jones, and Schoer, 1963, p. 97) The writing ability of the students was tested by using the Hudelson Typical Composition Ability Scale, "which required the reproduction of a narrative ('Slatter's Hill Snowball Fight')." (p. 98) The instruments used for the pre-test purposes were assumedly all well-designed tests, available at that time. RA did not give detailed specifications of these tests in terms of the author, the rate of validity or the rate of reliability, or any critique about the technicalities of these tests. Concerning the use of readers, the report mentioned the fact that "no data were offered on reader reliability." (p. 98) If the tests employed so far were all "objective tests," and used the multiple choice format or other similar formats, the issue of "reader reliability" would not be very significant, the only necessary item to verify being the counting of scores. On the other hand, if judgement was needed in deciding the

score for a certain test item, the reader reliability would be indispensable. Concern was actually directed toward the rating of the written composition, where five raters were employed, none of whom "knew which papers were written by which class." (p. 98) In this case, the technique of taking the average score replaced the reader reliability as was widely adopted in sport competition. For the post-test, ten objective tests were given in capitalization, punctuation, spelling, and grammar. Also, another measure was reported in RA, the mean total points for the approximately 30 papers per pupil written during the letter writing unit of the English class. The investigator was the only rater of these papers, and the mean total points of all the papers were used instead of the mean total points of the last few assignments, two facts which, according to Braddock, Lloyd-Jones, and Schoer, largely discounted the significance of her finding that the small class was distinctly superior in this measure. (p. 102) RA asserted that the most valid measure of composition ability for the post-test in this study was the Hudelson English Composition Scale (original narrative on "The Most Exciting Ride I Ever Had"). (p. 102) Again, the reader reliability was not offered. On top of these tests, a questionnaire was used six months after the start of the experiment to measure the attitude toward the small class and large class instruction. Nine students were reported to have been absent for the filling of the

questionnaires, five from the large class and four from the small class, though absences from other testing and instructional situations were not reported. Classroom observation was used to investigate the procedures of the actual classroom activities, and no specific data were reported except the fact that three observers, all of whom were university seniors training to be English teachers, came to sit in each class daily every three weeks and record pupil participation in the room, the time consumed in movements to and from the blackboard or in small-group activities, and the time devoted to each section of the lesson in each class (p. 98). Only the instructions to the observers were reproduced in the RAS, but not the actual observation reports.

To sum up, the instruments used in this study for measurements were of a variform nature. Well designed objective tests, composition scales, questionnaires, and observations were used as instruments to measure the beginning, middle, and end situations of the study. Although the application and accuracy of the instruments and the validity and reliability of the instruments and measurement still waited to be refined, the total consideration for the use of various instruments in this study for different types of measurements to be tested was certainly a strong point in this study.

Compared with the Smith study, the Kincaid study used carefully selected raters as instruments to measure the composition themes. The investigator examined "the records of ratings by more than forty staff members of the Department of Written and Spoken English," and selected nine staff members who "demonstrated a high degree of sensitiveness and reliability." Three raters selected from the group of nine were "willing and able to devote the time necessary" to rate each of the 320 papers in this study. This study employed "the usual rating scale for the Written and Spoken English course at Michigan State." A one (unsatisfactory) to ten (superior) scale was used in actual grading. Five equally weighted categories were graded on this scale: conventions of grammar, sentence structure, diction, organization, and content (p. 86). Several techniques were employed to achieve better rater reliability. First, to avoid variation in rater efficiency from day to day, all three graders judged all four papers of each student during the same rating period. To avoid rater's bias because of knowledge of who wrote a specific paper, the papers were made anonymous by using numbers assigned to each paper. To avoid rater's bias because of knowledge of how other raters rated the papers, the investigator made sure that no rater "had any indication of a score assigned by a previous rater." To avoid rater's bias because of knowledge of the sequence of the papers they

rated, the papers by the same student were separated by at least two other papers. RA cautiously pointed out that RAK did not state whether a grader could tell, at the time of rating, what score another grader had assigned the paper; nor whether the raters knew any or many of the student writers or could identify the papers of a particular group. The actual paper score was decided by averaging, out of three rater scores, the closest two scores, or all the three scores if they were equidistant or if no two ratings came within five points of each other. However, the rater reliability was not reported (pp. 86-87).

The Buxton study used three pre-tests: the Cooperative English Test A, Mechanics of Expression, Form 2; the Cooperative English Test B2, Effectiveness of Expression, Form 2; and an original essay written during a 50-minute period. Of the first two objective tests, parallel forms, Form X of each of the Cooperative tests, were used for post-tests (p. 62). To achieve satisfactory rater reliability in rating the two original compositions, the Buxton study used the following precautions: the anonymity of the students was preserved by thoroughly mixing the 257 examination booklets by one faculty member; a number from 1-257 was written on the cover and the first inside page; the cover page with student's name and number was removed; two raters both with more than twelve years' experience in teaching English in high school were used; a score sheet was prepared and used

which included fifteen categories; the two raters practiced with this score sheet on trial themes before using the sheet in actual scoring. RA writers cautiously pointed out that the pre-test papers and post-test papers were scored separately because of the fact that the papers were numbered from 1 to 257 only and not from 1 to 514, thus making the rating "useful for comparing the relative composition quality of the three groups but not for measuring progress from pre-test to post-test." (p. 65) On the score sheet, each student was given a basic mark of 70, and a maximum score was indicated for each of the fifteen categories, making a theoretical total of 300 (Table 3).

The above mentioned techniques together with a number of operational procedures, including rating every 25th paper together to insure rater consistency without changing scores given to previous papers, marking mechanical and grammatical errors in the margins and finding out the "error rate" (number of errors divided by number of words written and then multiplied by 1,000), and calculating the final score by averaging the scores made by the two raters to be divided by three to yield a scaled score based on 100, which facilitated comparison of the essay and objective test scores. All these precautions and procedures taken by the investigator yielded rater reliability coefficients estimated by the Pearson product-moment formula to be .91 for the pre-test themes and .88 for the post-test themes,

Table 3
SCORE SHEET

A. MARKS AWARDED:	Max.	Student
1. Basic Mark (Give every paper):	75	75
2. Material:		
a. Significance:	30	
b. Evidence of Critical Thinking (defining terms; recognizing generalizations; providing evidence; etc.):	30	
c. Originality (use of humor; exaggeration for effect; mock- seriousness; anticlimax; understatement; pretentious language used for effect; etc.):	30	
3. Organization:		
a. Title (interest and appropriateness):	5	
b. Introduction:	10	
c. Logical sequence of paragraphs:	10	
d. Unity within Paragraphs:	10	
e. Transition within paragraphs:	10	
f. General coherence:	10	
g. Effective conclusion:	10	
4. Sentences:		
a. Variety in sentence structure:	20	
b. General fluency:	10	
5. Diction:		
a. Exactness and vividness of nouns, verbs, adjectives, etc.:	20	
b. Interesting and appropriate figures of speech, comparisons, illustrations, allusions, quotations:	20	
	300	

Source: Buxton, in Braddock, Lloyd-Jones and Schoer, 1963, p. 64.

indicating a high degree of consistency between the two raters (Braddock, Lloyd-Jones, and Schoer, 1963, pp. 65-66).

RABK used a 45-minute objective test in English composition and a 450-word minimum, two-hour expository theme for both the pre-test and the post-test. One "experienced member of the Communication Skills staff" rated the pre-test themes, whereas three members rated the 10:3 class post-test themes, and two members rated the 10:2 post-test themes with another reader added in cases of marked discrepancy. It was not clear what kind of discrepancy was referred to here. In terms of reliability of the rating, the writers of RA learned from the writer of RABK that "the reliability of the rating was .83." However, it was not formally reported in either of the reports. Therefore, two concerns arose here. First, was the reliability of rating calculated among the raters? It was not clear from RA whether the raters were the same raters. Neither was it clear whether the raters of the post-test themes were "experienced." Secondly, the students chose one out of four topics on both occasions, for the pre-test and the post-test. Therefore, it was not clear whether the reliability of rating was calculated also based on the difference of topics. To expand the understanding of the writing instructional methods involved in this study, the investigator also administered several attitude tests to the students and instructors. A "Study of Educational

Attitudes" test was administered to the students both at the beginning and at the conclusion of the experiment with 66 multiple-choice items concerning the students' feelings about the importance of various communication skills or concerning their personal information and 12 more items concerning other college subjects interspersed among the 66 items. Besides, two versions of a questionnaire were administered separately to the students and the instructors near the close of the course, inquiring how they felt about the three methods under study (p. 112). Braddock, Lloyd-Jones and Schoer mentioned that there were some "unsatisfactory aspects" in this study which could have been avoided if a pilot study had been introduced (p. 107). Although the "unsatisfactory aspects" were not defined in RA, some improvements might have been suggested in the preparation of instruments for measurement. For example, more consideration could be given to how to design carefully the application of raters for the purpose of this study, and how to design questionnaires that tried to extract information from students and instructors who assumedly did not have the experience of the instructional methods except the one they were involved in to respond validly to the other two kinds of instructional methods.

RAH used a three-month pilot study on two classes before the main experiment. Two measures were administered to each of the two groups before and after the pilot study.

One was a short-answer grammar test, which was also used in the main experiment without any recorded difference. The other was a descriptive composition. The topic for the pre-test was "A Day in the Country," and the topic for the post-test was "A Day at the Seaside." (p. 73) The composition was to be completed in two 40-minute class periods each time. The measuring instrument for the student themes was a number of frequency counts identified to have reflected "the rate of growth of a child's maturing style." (Table 4)

The result of the instruments used in the pilot study was encouraging. The Formal Grammar Test, named Test C, "not only achieved a coefficient of reliability of $.923 \pm .014$ by the split-half method, but it yielded regular increases in scores with age (Table 5). The pilot study also revealed a series of correlation coefficients for the listed criteria (Table 6).

RBO used the California Test of Mental Maturity (IQ scores, mean 100, SD 16) and five pre-treatment compositions as the pre-test. The result of the IQ score was used as one of the criteria in matching the students of the experimental group and the students in the control group. The measurement used on the pre-treatment compositions was the score on words per T-unit calculated from the first ten T-units in each of the five pre-treatment compositions the students wrote. Six factors of syntactic maturity were identified and used in this measurement, namely, 1. words

Table 4
CRITERIA FOR MATURING STYLE

-
- a. Average length of the correct simple sentence.
 - b. Instances of the omission of the full stop at the end of a sentence and of inadequate separation between clauses, scored by subtracting the number of omissions from the number of sentences in which the stop was correctly inserted.
 - c. Number of words per common error. (Including the note at the foot.)
 - d. Number of different sentence patterns used by each writer, counting only correct sentences by accounting, for example, different types of adverbial clauses or inversion of the order of similar clauses as sufficient to provide a new pattern.
 - e. Number of non-simple sentences, whether correct or incorrect, from which the number of simple sentences was subtracted.
 - f. Number of subordinate clauses used--the actual number (not in proportion to the number of main clauses).
 - g. Total number of words written in a limited time.
 - h. Number of correct simple sentences, minus those containing an error (even though minus scores were frequent at 10 and 11 years of age).
 - i. Number of correct simple sentences containing two or more qualifying phrases, each phrase consisting of three or more words.
 - j. Total number of correct sentences minus the number of incorrect.
 - k. Number of adjectival clauses and phrases, each phrase consisting of three or more words.
-

(Adapted from Harris, in Braddock, Lloyd-Jones and Schoer, p. 75, 1963.)

Table 5
VALIDATION DATA FOR TEST C

Age Group	Average Score	Range of Score
11 years	8.1	0-16
12 "	13.0	4-22
13 "	22.8	12-34
14 "	33.1	19-48

(Source: Harris in Braddock, Lloyd-Jones and Schoer, p. 74, 1963.)

Table 6
RELIABILITY OF THE CRITERIA

Criteria	r
a. simple sentence length	.239
b. omission of full stop	.574
c. words per common error	.844
d. number of sentence patterns	.791
e. non-simple sentences	.393
f. number of subordinate clauses	.705
g. total words	.232
h. complex sentences	.685
i. simple sentences with two or more phrases	.447
j. total correct sentences	.839
k. adjectival clauses and phrases	.446

(Adapted from Harris, in Braddock, Lloyd-Jones and Schoer, p. 76, 1963.)

per T-unit, 2. clauses per T-unit, 3. words per clause, 4. noun clauses per 100 T-units, 5. adverb clauses per 100 T-units, and 6. adjective clauses per 100 T-units. The post-test was also five post-treatment compositions measured in the same T-unit score and against the six factors of syntactic maturity. An ordinal evaluation was made of the post-treatment compositions by eight outside evaluators. As O'Hare declared that the unreliability of composition ratings was "notorious," this study used the system of forced choices between matched pairs and used the ordinal data concerning the students' written compositions. Eight evaluators were instructed simply to pick the better composition from a pair of papers based on their decision equally on ideas, organization, style, vocabulary, and

sentence structure (O'Hare, 1971, p. 50).

RF used three rough drafts as the pre-test and three final drafts as the post-test, focusing on the effect of different types of editing instructions between the drafts. In the six categories of focus, sequence, support, sentences, language, and flavor of both the pre-test papers and the post-test papers and in the category of degree of change between the pre-test and post-test, all the student papers were rated on a 1-5 scale by three judges, the reliability of whose ratings was listed in Table 7.

Table 7
RATER RELIABILITY

Category	Pre-test r	Post-test r
focus	.87	.90
sequence	.86	.90
support	.90	.91
sentences	.91	.92
language	.90	.90
flavor	.89	.90
degree of change (pre-post)	.84	

(Adapted from Beach, 1979, p. 115.)

When these instruments and measurements were carefully examined, a few points should be brought to attention. First of all, carefully designed objective tests, though they had high documented validity and reliability, were not appropriate for holistic composition quality rating. Each objective test was designed only for each specific purpose, that is, to measure what it was designed to measure:

punctuation, expression, grammar, etc. For overall composition quality rating, raters were used where rater reliability could be a critical issue. To increase rater reliability, training and randomization of test essay papers on an anonymous basis seemed to be the most important technique. Secondly, the reliability of using involved raters could be questionable because the instructional focuses could be more sensitized in rating than using outside raters, and there was the probability of the raters knowing some or many of the students merely by their handwriting style and bringing in rating bias based on previous knowledge about the subjects. Thirdly, the approaches of quantification might be used, but it must be done with caution, as quantified data tended to be one-dimensional, whereas holistic rating involved human judgement based on various subhierarchical structures.

The statistic procedures used in these studies were varied. RAS used range, mean scores and standard deviations to justify the pairing of the students of the small class with selected students of the large class; in reporting, RAS also used percentage figures to display all types of measurements. Except those statistics, no test of homogeneity or test for significance was discernible in RAS. It was not clear, therefore, whether the differences observed in this experimental study were by mere chance distribution or were significant at a certain confidence

level.

RAK used mean scores and an ANOVA to test the homogeneity of the grouped students for pre-test. RAK also used the means scores and the ANOVA to seek significant differences among the mean scores of the different groups of subjects by topic and examination pressure and also to seek significant differences among the mean scores of combinations of the groups of subjects by examination pressure. Further on, RAK used Ebel's "between-rater" variance test (1951) and found the "between-rater" reliability to be .83 (Braddock, Lloyd-Jones, and Schoer, 1963, p. 90). After that, RAK used a procedure described in detail on pages 75-76 of the study, but not described in RA, to determine how much greater than "theoretical differences" the difference actually measured between two papers had to be in order to be significant among different topics. When the four different pairings of papers were examined, more variations than theoretically expected were found. In addition, a rank order was obtained of the students' writing ability based on the highest score among the four papers, and another rank order was obtained based of the average score of the four papers. When the two kinds of ranking were compared, more variation was found in the upper 25% of the students than the lower 25% of the students. Finally, the chi-square test was applied on a series of variables by using the frequency count technique to find out whether

content factors, or timing factors, or the factors of similarity or dissimilarity of topics, or the factors of examination pressure, either among the upper 25% or 50% or among the lower 25% or 50% of the students, were responsible for the variations observed.

RABX tested the degree of homogeneity of the variances among the three groups of subjects before it conducted ANOVA and ANCOVA with the essay test scores. Then the differences between the post-test means was adjusted by an ANCOVA in terms of differences in the pre-test means. Significance was determined by applying the t-test to the differences between the adjusted post-test means. When the tests of homogeneity of the within groups variances in the essay pre-test found the ratio of the control group to the revision group to be significant at the .05 level, Bartlett's test was applied and the yielded chi-square value of 2.818 for two degrees of freedom justified the use of ANOVA and ANCOVA. When ANCOVA yielded a significant F-ratio of 4.98 with 2/251 degrees of freedom, the t-test was applied to the differences between pairs of adjusted post-test means. Furthermore, RABX used the Garrett's method to test the within-groups homogeneity on the means of the students' final examination scores in all their subjects and also tested the significance of the differences of the means in order to calculate the effect of the instructional factors. Finally, an itemized analysis was made of the gain in mean

raw scores from pre-test to post-test by each group on each of the fifteen categories on the score sheet.

RABK was reported to have used "careful statistical procedures" in RA, just because of which it was suggested that, even if the results were not statistically significant, the results might still be meaningful (p. 107). First of all, RABK used the ANOVA to test the homogeneity of the classes on each of the pre-tests and placement test composite score. Then a groups-within-treatments ANCOVA was designed to test the differences among methods of instruction by using individual scores. A test was first conducted to determine whether there were extreme or systematic differences among instructors before the ANCOVA. The differences observed among methods of instruction were then tested by the ANCOVA (pp. 112-113). After that, RABK used an ANOVA utilizing a treatment-by-level design (four levels of approximately equal sizes of students based on the composite score) to test whether students of different levels of academic aptitude achieved more or less in theme writing under each method. In the end, an ANOVA was also applied to the "Study of Educational Attitudes" test.

RAH used a series of t-tests to test the differences between means of the formal grammar and direct method groups at each school for each of the eleven criteria. RAH also conducted a number of correlational computations on the scores collected to determine the correlational

coefficients. It should be pointed out that the level of confidence in this study was not defined in terms of percentage. Instead, RAH used "V" (very reliable), "F" (fairly reliable), and "U" (unreliable) to indicate the test results.

RBO began by testing the assumption of homogeneity of the pre-treatment mean scores on the six factors of syntactic maturity with the t-test for two independent samples assuming unequal variance, a method described by Dixon and Massey (1969, p. 119). Then separate t-tests were run on the difference between pre-treatment and post-treatment scores to determine whether the differences observed were significant for either the control group or the experimental group. After that, t-tests of post-treatment mean scores of the two groups on the six factors, t-tests of the pre-post change scores of the two groups for two independent samples assuming unequal variances, separate t-tests of the post-treatment mean scores by teachers for the control group and for the experimental group, t-tests of post-treatment mean scores by male and by female students in the control group and in the experimental group, correlational coefficients or multiple correlational coefficients between 1. pre-test scores and post-test scores, 2. pre-test and IQ scores and post-test scores, 3. IQ scores and post-test scores, and 4. pre-post change and IQ scores and post-test scores for the experimental group,

multiple regression of post-test score on IQ and pre-test score on the six factors for experimental group, and chi-square tests on differences observed versus expected of the forced pair evaluation of the eight raters, chi-square tests for comparison between two types of writing, and chi-square tests for comparison among the eight teacher-evaluators were in turn conducted for statistical treatment.

RF first used ANOVA developed by Godshalk, Swineford, and Coffman (1966) to test the inter-judge reliability for the three judges' ratings for all drafts on the degree-of-change and the quality ratings (Beach, 1979, p. 115). Separate one-way ANOVA's were run in order to determine whether the sessions in which the students wrote over a three-week period had any effect on subjects, degree-of-change scores, or the quality ratings of both rough and final drafts across the three topics. A $2 \times 2 \times 3 \times 3$ mixed factorial ANCOVA of the means of three judges' ratings for each version was conducted for each dependent measure with S's nested in grade level by sex by treatment combinations and crossed with the repeated measure, topic, with rough draft ratings and writing apprehension scores as covariates. Where F tests were significant, Scheffé procedure was employed to determine significant comparisons between adjusted cell means (pp. 115-116).

It should be borne in mind that the studies selected for analysis in this study were carried out in a time span

of more than sixty years. The increasingly sophisticated patterns within the chronological frame indicated that the statistical techniques were getting mature along with the cumulative knowledge base. The progress of comprehension and insight were apparently of a chronological nature. This temporal dimension, therefore, can be best described in this composite hierarchy theory. Under normal situations, a higher level of sophistication can be detected in a study conducted at a later time. Besides the temporal factor, the factor of purpose of each individual experiment also decided the level of sophistication of the statistical procedures to be adopted. A third factor which would affect the choice of statistical procedures was the type of data collected from the experiment as dictated by the design of the study.

Inquiry takes different forms and different types of research projects answer different types of research questions. An "investigation" might be useful only in suggesting issues for further research because of its lack of rigor in controlling or sampling. A "survey" might be useful only in detecting possible cause-effect relationships because of its lack of rigor in controlling. A "study" might be useful only in defining the properties of the group of subjects under study because of its lack of rigor in sampling. Only the "experiment" can provide insight of the cause-effect relationship by means of the rigorous sampling and controlling procedures adopted (Guba, 1963, pp. 242-

248).

Conclusions drawn from the quantitative data gleaned from the experiments posed another aspect of concern for research specialists. If the conclusions of a study were to be valid, the design of this study had to be protected against a number of sources of external invalidity, such as pre-test-treatment interaction, multiple-treatment interference, selection-treatment interaction, specificity of variables, experimenter effects, and reactive arrangements (Bracht and Glass, 1968).

When the conclusions of the seven experimental studies were examined, a few concerns were brought up. For example, in RAS, Smith did not consider the Hawthorne effect in her study. As she was the sole experimenter, she might have brought into her class a tendency to surpass the usual efforts because of heightened interest in the experiment. Her conclusion that "a school could save 28 percent of the teacher expense if it assigned teachers three classes of 50 pupils each and clerical assistance for three hours a day rather than five classes of 20 pupils each and no clerical assistance" was met with doubt, as she herself asked the question "how many classes of 50 a teacher may teach in one day without harm to herself or to the pupils." (Braddock, Lloyd-Jones, and Schoer, 1963, p. 106) If we examine what her study told us, we could safely conclude that the experience and dedication of an excellent teacher like Dora

Smith could give a class of 50 students that kind of instruction so that their learning in composition, and in other aspects of English study, could match the learning of students in a smaller class of 21. It remained to be answered whether any teacher could achieve the same results.

The Kincaid study concluded as follows:

1. Where a group of college freshmen is concerned, the average quality does not seem to be affected by variations in the day-to-day efficiency of individuals or by the pressure of the examination situation.
 2. Variation from expository to argumentative mode of discourse does not seem to affect the average quality of the writing of a group of freshmen.
 3. Variations in the quality of writing of individual freshmen did seem to result from variations in the day-to-day efficiency of students, especially of the better writers, but the pressure of final examination conditions did not appear to affect adversely the writing of either the better or the worse writers.
 4. Although the writing of the better writers did not seem to vary significantly in quality according to whether the students were writing expository or argumentative papers, the quality of writing by the worse writers did reveal a significant number of variations according to mode of discourse.
- (Braddock, Lloyd-Jones, and Schoer, 1963, pp. 94-95).

In this study, the conclusions drawn from the statistical treatments were well within the scope of analysis. All the conclusions were carefully phrased so that the conclusions were not over-generalized.

The Buxton study concluded as follows:

1. Raters who use a scoring system which they understand and endorse, who have practiced rating papers by this system, and who make periodic

checks on their rating may achieve a high degree of consistency in rating essays, or themes.

2. Since the students in the three groups did not differ significantly in their gain on the two objective post-tests but did differ significantly in their gain on the post-test essay examination, it may be seen that essay examinations can be used to measure changes which may not be measured by objective tests.

3. College freshmen whose writing is graded and thoroughly marked and criticized and who revise their papers in the light of these matters can improve their writing more than college freshmen whose writing receives a few general suggestions but no grades or intensive marking and who do not revise their papers.

(pp. 69-70)

The Becker study concluded as follows:

1. There seemed to be no consistently significant difference in the effectiveness of the three methods in the two-semester regular course.

2. The Normal method may have been more effective than the Bibliography method for teaching theme writing to freshmen in the one-semester accelerated course, but the observed difference may well have been the result of chance.

3. It is possible that the Kinescope method would have fared better had lecturers more experienced with television made the kinescopes and that the Bibliography method would have fared better had the bibliographies been more selective.

4. It is also possible that both the experimental methods would have been more successful if the instructors had had more experience with them, but the attitudes of the instructors seem to work against the likelihood.

5. Although the Bibliography method permits the saving of classroom space and of instructor time in class, it may be less effective as a method of teaching composition to college freshmen who are generally of above average academic ability.

(pp. 115).

Braddock, Lloyd-Jones, and Schoer pointed out that if the Becker experiment had been preceded by a pilot experiment--permitting the development of more satisfactory kinescopes and bibliographies and perhaps a more supportive attitude on the part of the instructors, the kinescope and bibliography methods might have been shown significantly superior to the normal method. It should be pointed out that when the investigator began the study, the treatment had not been carefully designed. Both the investigator and the reviewers were in doubt as to what this study could really tell us. These speculations were not logically inferred from the statistical treatments, but from some preconceived, yet not quite materialized, speculations. For experimental studies, it was very important to give careful consideration to every detail in this study. As soon as the collection of data began, it would be too late to make any changes. The issue of the attitude of the instructors was also important. The overwhelming rejection of the bibliography method as was reflected in the questionnaire could have contributed to the effect of that specific method under study.

The Harris study concluded as follows:

1. The low achievement of the Formal Grammar group on Test C may be seen as a major factor throwing doubt on the advisability of studying formal grammar in the early part of the Secondary School.
2. The failure to profit from instruction in [traditional formal] grammar is thus not confined

to any one educational environment or category of children" (of those he studied).

3. It seems safe to infer that the study of English grammatical terminology had a negligible or even a relatively harmful effect upon the correctness of children's writing in the early part of the five Secondary Schools.

(p. 83)

It should be noted that in the conclusions of RAH, the term English grammatical terminology referred to the terminology of traditional grammar. It did not give any indication whether the use of structural grammar terminology or transformational-generative grammar terminology would be equally ineffective. It should also be noted that the results were obtained from students aged twelve to fourteen. The study did not indicate that the same results could be expected from older students. Because of the lack of randomization of the subjects under study, it would be unsafe to infer that the same results could be expected elsewhere.

The O'Hare study concluded as follows:

The experimental group wrote compositions which were syntactically different from the compositions written by the control group. The experimental group wrote significantly more clauses and these clauses proved to be significantly longer. As a consequence the experimental group wrote T-units which were significantly longer than those of the control group. When compared with the normative data presented by Hunt (1965), the experimental group's compositions showed evidence of a level of syntactic maturity well beyond that typical of eighth graders and in many respects quite similar to that of twelfth graders (O'Hare, 1971, p. 67).

It should be noted that in the conclusion of the O'Hare study, the conclusions were drawn directly from the statistical treatments. The terms of experimental group and control group were used instead of the more pervasive term "students." When O'Hare used the normative data for comparison, the reference was actually made in the conclusion, which displayed a high degree of observance to the rigor of experimental designs.

The Beach study concluded as follows:

Students who were provided between-draft teacher evaluation showed a greater degree of change, higher fluency and greater differences in support on final drafts than students employing guided self-evaluation forms and students receiving no evaluation (Beach, 1979, p. 119).

The conclusion in this study just stated the results of the experiments. However, it should be noted here that the students were the 103 students in three 10th grade and two 11th/12th grade classes at a high school located in a predominantly working-class neighborhood in Minneapolis in 1979. Any effort at inferring onto students of other characteristics should be with extended care.

Conclusions drawn from a study could be valid only if these conclusions were generalized at the right level of the hierarchy, if other requirements for validity and reliability were also met. These levels of the hierarchy structure could involve age, IQ, personal background and experiences of the students as well as those of the

instructors, school environment, and community support.

Strengths, Weaknesses and Applicability

There were countless books published discussing experimental research methodologies in the western world. A number of them were related to the experimental research methodologies used in humanities and educational situations, including works by Campbell and Stanley, 1963; Roscoe, 1975; Cohen, 1977; Nunnally, 1978; Ferguson, 1981; Gay, 1987; and numerous others.

The traditional perspective on the empirical study is based on a well defined set of serial steps as follows:

1. Theory construction
2. Deviation of theoretical hypotheses
3. Operationalization of concepts
4. Collection of empirical data
5. Empirical testing of hypotheses.

(Babbie, 1973, p. 227)

For all the concerns about the validity and reliability of experimental studies, this type of studies so far has offered the most substantial results in the research field of composition. Although the format of these studies still has involved issues for debate, and the conclusions from these studies are still regarded with suspicion, the achievements made by means of this empirical approach are above any doubt. This type of study is especially desirable when a certain detected cause-effect relationship needs to

be verified. With this approach, different forms of research designs are available for different purposes, different conditions, and different resources. Nevertheless, caution is highly recommended for the use of this type of study for various reasons.

Campbell and Stanley (1971) identified eight major threats to the internal validity of experimental study:

History: occurrence of any event which is not part of the experimental treatment but which may affect performance on the dependent variable

Maturation: physical or mental changes that may occur within the subjects over a period of time

Testing: improved scores on a post-test resulting from subjects having taken a pre-test

Instrumentation: unreliability, or lack of consistency, in measuring instruments which may result in an invalid assessment of performance

Statistical Regression: tendency of subjects selected on the basis of their extreme scores in pre-test to score less extreme scores in post-test

Differential Selection of Subjects: situation where the groups may be different before the study even begins, and this initial difference may at least partially account for post-test differences

Mortality, or attrition: fact that subjects who drop out of a group may share a characteristic such that their absence has a significant effect on the results of the study

Selection-Maturation Interaction, Etc.: initial advantage or profit for already formed groups because of maturation, and Etc.

refers to interaction of selection with other factors than maturation.

(Campbell and Stanley, 1971)

These threats to the internal validity should be carefully considered in designing experimental studies. Besides, threats to the external validity should also be taken care of because these threats would greatly limit the generalizability of the study. Bracht and Glass (1968) define the six threats as follows:

Pre-test-Treatment Interaction: difference in subjects' responding or reacting to treatment because of pre-test

Multiple-Treatment Interference: carry-over effect from an earlier treatment which makes it difficult to assess the effectiveness of a later treatment when the subjects have received more than one treatment in succession

Selection-Treatment Interaction: similar to differential selection of subjects, difference that occurs when subjects are not randomly selected for treatments

Specificity of Variables: fact that a given study is conducted (1) with a specific kind of subject; (2) based on a particular, operational definition of the independent variable; (3) using specific measuring instruments; (4) at a specific time; and (5) under a specific set of circumstances, which threatens the generalizability of the study

Experimenter Effects: effects unintentionally brought into experiment by the researcher, passively or actively (Rosenthal, 1966)

Reactive Arrangements: a number of factors associated with the way in which a study is conducted and the feelings and

attitudes of the subjects involved.

Desirably, empirical studies should be limited in social and behavioral sciences first to lower levels of complexity in the theoretical hierarchy structure. The higher the level of complexity was, the less valid the control could be for the empirical study (Guba, 1963, p. 240), making the conclusions drawn based on the less valid control also less valid. By nature, empirical study should be designed to focus on variables singled out by means of controls and to allow research problems to be clearly defined and examined by commonly accepted and standardized techniques and methods so that straightforward answers to problems studied under experimental conditions are obtained in a logical and sequential fashion, building upon each other (Wolf, 1986, p. 9). If the level of complexity was very high, too many variables, most of them being confounding variables (Guba, 1963, p. 239), could be buried under the highly generalized terms, and even the artificial and usually distorted definitions would not be able to clarify the concepts under study. Therefore, for the purpose of policy making, the results of such studies tended to be too muddled, unproductive, and unscientific (Wolf, 1986, p. 10).

The type of data was critical in deciding which statistical procedures to adopt for each individual study. The property of the data, the measuring instruments from

which the data were collected, and the relevance between the research questions and the data collected were among the most crucial issues for the validity of the study.

The theory of hierarchy looked to the lower levels for mechanism and that is where the empirical study should really begin. The task before the researcher should be to break down the concepts of research interest to the lowest possible level of complexity by means of the hierarchy structure and, at the same time, maintain the integrity of the objective of the research in view in order to examine the effects of those variables in all their phases. When these basic variables have offered a sound base for comprehension, a combination of variables could be studied for their composite effects in different dimensions.

In conclusion, this chapter discussed the several experimental studies selected, putting emphasis on the theoretical structures of these studies, the components of these studies, and the technical aspects of these studies, and discussed some issues involved in methodological concerns in these studies that could have been improved. In order for the discussion to be beneficial to Chinese professionals, a number of suggestions were also provided for the construction of guidelines for Chinese professionals in conducting experimental studies in teaching English composition in China.

CHAPTER IV
DESCRIPTIVE RESEARCH IN TEACHING WRITING
Philosophies, Theories and Concepts

Frequently, the purpose of a study is not to identify some cause-effect relationships as constantly seen in the case of experimental studies. The purpose can be to identify the characteristics of a group of subjects of interest. The purpose can be to examine the statistical profile of a targeted population. The purpose can be to examine the accuracy of a research instrument. The purpose can be to identify a correlation between two variables. The purpose can be to explore an unknown area. For these purposes, researchers tend to adopt descriptive research techniques for their desired results.

It can be seen from Table 1 that, of the twenty reports selected for this study, a number of research reports belong to the category of descriptive research. Emig's case study RBE (1971), Donlan's survey study RC (1974), Mellon's measurement study RDM (1975), Stewart's measurement study REC (1978), Applebee's observation study RHA (1981), Grobe's measurement study RHGC (1981), McCulley's measurement study RK (1985), Swanson-Owen's case study RL (1986), Chesky's measurement study RMC (1987), and McCann's measurement study RO (1989) are all cases of descriptive studies. The differences between the descriptive research and experimental research lie in the fact that descriptive

research does not use treatment variables in the research.

The theoretical frameworks, conceptual systems and philosophical orientations of the research studies are inextricably bound together throughout all phases of the research studies. In reality, the inquiry process, whether it is scientific or otherwise, occurs within the context of researchers' personal experiences, general sociocultural frameworks, and philosophical traditions, including their assumptions about the nature of reality, knowledge, and values (Bredo and Feinberg, 1982). These factors also interact with the social science constructs which the researchers take as their reference points for inquiry. In consequence, research methods and methodological choices are integral to the theoretical systems brought to the inquiry process by the researchers (Goetz and LeCompte, 1984).

The theoretical basis for descriptive studies was somewhat different from the theoretical basis for experimental studies. According to Guba (1963), inquiry was classified into four types based on the degree of observance of the criteria for external validity and the criteria for internal validity. When the criteria for both internal validity and external validity are met, the study can be regarded as an experimental study. Otherwise, the study can be only either a study, a survey, or an investigation (Table 8).

Table 8
FOUR TYPES OF INQUIRY

Presence of		Type
S*	C*	
-	-	Investigation
+	-	Survey
-	+	Study
+	+	Experiment

*S (sampling) indicates meeting criteria of external validity; *C (control) indicated meeting criteria of internal validity. Adapted from Guba in Culbertson and Hencley (1963, p. 242).

In most cases, the observance of the criteria is achieved to a certain degree, but not to satisfaction. These studies will still be regarded as having observed the criteria though with weaknesses. Descriptive studies do not rigorously demand strict observance of criteria for both internal validity and external validity. That is why the data gleaned from descriptive studies should only be regarded as valid descriptive data. For inferential purposes, great caution has to be exercised on the use of descriptive data.

In practice, different types of studies serve different purposes. As long as the study successfully answers the research questions defined, the type of research will be regarded as a good research type.

Descriptive studies included survey studies,

observation studies, measurement studies, and case studies. Although all these studies fall under the category of descriptive studies, they certainly have their own distinctive features. Survey studies are characterized by the use of questionnaires distributed to a group of selected subjects for written responses. Observation studies are characterized by classroom observations over an extended period of time. Measurement studies are characterized by the adoption of carefully designed measuring instruments for repeated measurements of the behaviors of subjects. Case studies are characterized by a carefully selected small group of subjects for in-depth observation and measurements over an extended period of time.

Frequently, different types of research techniques are combined in a single study. With some experimental studies discussed in the previous chapter, survey techniques were also adopted to gather data for related research purposes.

Survey, Observation, and Case Studies

Of the descriptive studies selected for analysis in this study, Donlan's study (1974) was an example of survey research. Applebee's study (1981) was an example of observation research. Emig's study (1971) and Swanson-Owen's study (1986) were examples of case studies.

Graves and Murray's study (1980) was a unique example

of descriptive research in the form of journal writing. The researcher and the writer cooperatively put down their thoughts and sentiments captured during their working processes. The writer was doing prewriting, writing, editing, revising, and getting published on a three-hundred-page novel between September 19, 1979, to December 8, 1979. The researcher was working on an NIE-funded research project on the group writing behavior of sixteen children of 6 to 9 years of age, doing videotaping, hand-recording, note-transcribing, and data-analyzing during the two successive years. The researcher and the writer offered a large amount of insightful information in this report in the form of their journals.

Some of the journals they wrote read as follows:

Murray

Steph pointed out in my draft of a chapter of the Donovan-McClelland book that I incorrectly used "for" in the second paragraph. She suggested "because." My reaction was normal; I rejected her suggestion, as I would reject any editor's suggestion. I overreacted and rewrote the whole paragraph. When I receive criticism, I normally put the draft aside and start a new one. It is probably the way I reestablished control over my territory. Childish. But the paragraph was better. "For" became "who."

Graves

Our data show that children as young as eight years of age are capable of writing to find out what they mean. For such children, six to ten unassigned drafts are not unusual.

Murray

Revision of one article I am doing is not revision by my definition. There are no new visions or insights, just simple editing for

clarification. It is a bad article. It needs no work, has no possibilities hidden between its words.

Graves

The "can I have another piece of paper" syndrome is in many classrooms, especially where there are good readers who write infrequently. Their writing tools lag well behind their ability to read. They are painfully aware of the discrepancy between their written text and what they wish they could say.

Murray

It took me from spring to autumn to create the new beginning of the novel. To put it differently, it took me months to hear what the novel had to tell me about its story. If I had listened to the draft, it would have told me how to begin the novel. It did tell me. When I found the right beginning there were no major changes to be made in the text. The novel was waiting to be begun that way.

Graves

Students need help with the process of gaining distance. Teachers help through the writing conference. "What did you have in mind here? Underline the one line that says more than any other what this selection is about."

Murray

Much of the reordering in the text is making sure that the most important material is at the point of emphasis in the paragraph, in the

sentences, in the scene, where are the points of emphasis? At the end and at the beginning. The important information must be those points.

Graves

The work can take over when the teacher consciously works for students to find their own voices and to be responsive to the effect of voice on information. Students must teach teachers about their subjects, whether it be grade one or a dissertation.

(Graves and Murray, 1980)

Donlan (1974) did his survey study as a pilot study to

discover some answers to the question of how important writing was. Donlan found it necessary to determine how much difference there was between what research literature suggested and what the teacher in classroom practiced. To achieve this purpose, besides a literature review of some of the research reports and summaries, he conducted a survey among Riverside area teachers of kindergarten through grade 12. The research questions raised for this survey study were as follows:

- (1) How much writing is assigned in a given year by selected teachers from different content areas, kindergarten through grade twelve?
- (2) What is the nature of the assignments?
- (3) What teaching techniques are used with the assignment of writing?
- (4) Who is perceived as responsible for the teaching of writing?

Donlan designed a questionnaire (Appendix B) focused upon the above research questions and asked each of the thirty students in his class to interview about five teachers in the Riverside County. Some students interviewed more and some less. A total of 123 teachers in the area were interviewed, from kindergarten through graduate school. Twenty-seven questionnaires from college and university teachers were eliminated from the present study for another study. Four questionnaires came back too late for the tabulation. Therefore ninety-two responses from teachers of kindergarten through grade twelve in the Riverside area were

included in this survey study. When the tabulation of the information was completed, Donlan presented the results in his report RC (Donlan, 1974, pp. 254-256).

In discussing the efficacy of the survey research design, Donlan admitted toward the end of his report that "A survey has limitations. First of all, a survey rests on the assumption that responses are indicative of truth." (p. 262) It is true that a survey has limitations, as well as all other kinds of research designs. As to the claim that a survey rests on the assumption that responses are indicative of truth, it might not be very accurate. "Truth" as used in this context, could refer to the factual truth as well as to the attitudinal truth. Some of the items in the survey questionnaire could extract factual information to a substantial degree, while others could extract a certain amount of attitudinal information, or information revealing how the respondents viewed these question items. It is important that these two types of information be distinguished in interpretation. In actual social science research practice, survey research has been looked upon as one research approach with a very great social, scientific, and educational potential (Babbie, 1973, p. 372).

Based on the tabulations of the responses recorded in the ninety-two questionnaires, Donlan formulated the following eleven hypotheses concerning the practice in and attitude toward writing instruction of teachers in the

Riverside area:

- (1) There is a negative correlation between the length of a writing assignment and the teacher's tendency to assign it; however, the tendency toward longer assignments increases as the grade level increases.
- (2) Regardless of grade level or content area, teachers tend to favor frequent writing assignments. The tendency decreases as the grade level increases and tends to be stronger in the areas of social studies and language arts than in science and mathematics.
- (3) Exposition and reporting tend to be the most popularly assigned forms of writing. However, narration is the most popular form of language arts teachers, particularly those in elementary schools.
- (4) Most teachers, regardless of grade level or content area, tend to assign writing as an extension of classroom experience, rather than as a summary of such experience or as a substitute for such experience.
- (5) Teachers tend to favor writing in class as the most popular method of teaching writing; they also tend to feel the responsibility of teaching writing should not be shared with the students, as they did not use student editing committees.
- (6) Teachers, regardless of grade level or content area, tend to avoid indicating and correcting faulty sentences as a method of correcting papers.
- (7) Teachers tend to make more comments in the margins of papers than at the end, although this may vary slightly with the content area.
- (8) Teachers tend to evaluate papers on both form and content, placing more emphasis on content.
- (9) Teachers tend to grade papers with a combination of letter grade coupled with comments. The use of comments-only, popular among primary teachers, tends to decline as the grade level increases. In addition,

comments-only appears to be more popular among language arts and social studies teachers rather than among mathematics and science teachers.

- (10) Teachers tend to favor student files as a method of disposing of returned papers. Of the content areas, only language arts appears to deal with papers in any systematic way.
- (11) Teachers, regardless of grade level or content area, tend to feel the responsibility of teaching writing should be that of the content area teacher. However, a lesser percentage of English teachers feel the teaching of writing is solely their responsibility than the percentage of content area teachers who feel the English teacher is responsible for the teaching of writing.

(Donlan, 1974, pp. 252-254)

In RC, Donlan did not indicate whether the teachers interviewed were randomly selected. Nor did he indicate whether there were any subjects who did not want to be included in this study when requested. Based on the descriptive survey data, Donlan generated a number of hypotheses for future study. Therefore, the only potential contention regarding this process is whether the directions suggested in the hypotheses were reasonable. In other words, in using these hypotheses for further studies, would it be misleading to follow the directions of the hypotheses identified in RC? Or would it be more reasonable to design the hypotheses with no specific direction at all?

Applebee's observation study RHA (1981), a large-scale investigation, identified the following research questions:

- (1) What is the nature of the writing tasks currently being asked of secondary school students?
- (2) What are the teachers' purposes and techniques in making writing assignments?
- (3) What is the extent to which the characteristics of student writing vary with the subject area, grade level, and instructional goals?

In dealing with this list of highly complex questions, Applebee used two related strands of research, one of classroom observations of writing assignments and related instruction in two midwestern high schools over a full academic year, and one of a national questionnaire survey of teachers in six major subject areas: English, foreign language, mathematics, science, social science, and business education (Applebee, 1981, p. 9). The observation study that began in October of 1979 focused on the ninth and eleventh grade of both a laboratory school and a city high school. A total of 309 classes were observed, of which 259 classes fell into the main sampling frame (p. 14). The classes to be observed were randomized. A simple log with space to note time to the nearest minute (by a digital display) was adopted to record (1) types of class activities, (2) writing or note-taking activities, (3) time of change or interruption of activities, and (4) questions raised by teacher or students during whole-class discussions.

The questionnaire survey was conducted half a year

later in three steps. The first step was a trial study among university faculty. The second step was a pilot study among teachers of the observed schools. And the third step was a national study among teachers in six content areas mentioned before recommended by principals selected through stratified randomization in the whole nation. The survey instrument was largely self-coding, with a small number of open-ended questions addressing general attitudes. In addition, two copies of anonymous sample assignments done by their students were solicited from each participant teacher (p. 16). The response rate of this survey was .83, and the 754 usable questionnaires formed 68 percent of the total from the 1,108 recommended teachers.

From the coded activities of the 309 observations, a list of identified writing activities was generated. The classification of the activities still left room for discussion, but this overview accurately described the various types of writing activities that occurred in the classroom (Table 9).

Strictly speaking, a valid classification should be mutually exclusive between the categories. While the "use of writing" was the main concern in this classification, the first set of differences seemed to be in the form: the composing use and non-composing use of writing. The second set of differences seemed to be in the content: informational use or non-informational (imaginative) use of

Table 9
USES OF SCHOOL WRITING

Writing without composing

Multiple-choice exercises.

Fill-in-the-blank exercises (answered with less than a sentence).

Short-answer exercises (brief, one or two sentences per question).

Math calculations.

Transcription from written material (copying).

Transcription from oral sources (dictation).

Translation.

Other mechanical uses.

Informational uses of writing

Note-taking.

Record, of on-going experience. (This is what is happening.)

Report. Retrospective account of particular events or series of events. (This is what happened.)

Summary. Generalized narrative or description of a recurrent pattern of events or steps in a procedure. (This is what happens; this is the way it is done.)

Analysis. Generalization and classification related to a situation, problem, or theme, with logical or hierarchical relationships among generalizations implicit or explicit.

Theory. Building and defending at a theoretical level, including implicit or explicit recognition that there are alternative perspectives. Hypotheses and deductions from them.

Persuasive or regulative uses of writing. (Any instances in which the attempt to convince overrides other functions or in which rules are given and compliance assumed.)

Other informational uses.

Personal uses of writing

Journal of diary writing, for own use.

Personal letters or notes, where main purpose is "keeping in touch."

Other personal uses.

Imaginative uses of writing

Stories

Poems.

Play scripts.

Other imaginative uses.

Source: Applebee, 1981, p. 29.

writing. The third set of differences seemed to be in the purpose: use of writing for self or use of writing for others. In this list of use of writing, writing with composing and writing without composing, factual use of writing and non-factual use of writing, personal use of writing and non-personal use of writing, these seemed to be three different dimensions merged together, namely, form, content, and purpose. Therefore, a three-dimensional model can be set up for a dichotomous classification of classroom use of writing instead of the one-dimensional classification used here with much overlapping and sharing among the categories.

In the synopsis of this study, the researcher offered his definitions of good, better, and the best lesson as follows:

Good Lesson

there is an orderly variety of tasks for
students to perform
assignments are clear and their purpose
evident
students perform teacher-designed tasks
grades are used as the primary motivation
when assigned, writing is used as a measure
of student knowledge or performance
the predominant teaching technique is
teacher-led class discussion

Better Lesson

students are actively involved in teacher-
designed tasks
the teacher maintains a high level of student
interest
the teacher encourages a free flow of give
and take
the teacher incorporates student experiences
into the lesson

the predominant teaching technique is
teacher-led class discussion--but student-
led discussions are encouraged
students are prepared for writing assignments
with prewriting activities such as
audiovisual presentations or modeling
there is a climate of trust between teacher
and students

Best Lesson

students assume an active role in their own
learning
the teacher encourages students to explore
and discover and seldom dominates the class
students' own experiences are freely
incorporated into class discussions
students are enthusiastic about their work
writing is viewed as a means of learning and
emerges naturally out of other activities

(Applebee, 1981, pp. 104-105)

Case study is a unique type of research approach. A small sample is identified and different types of research skills are applied. Frequently, repeated measurements are taken in various forms over an extended period of time. The results from case studies usually reveal details in depth and suggest new direction for research. According to Heaphey (1963), case study was divided into two different approaches, one of preparation of case reports to capture history, and one of study of cases in terms of events and incidents to further the goals of social science as a nomothetic discipline (Heaphey, 1963, p. 276). Emig's study (1971) and Swanson-Owen's study (1986) were two case studies of events and incidents in order to further the goals of research in composition as a nomothetic discipline.

Louis M. Smith (1979) and Goetz and LeCompte (1984)

held that case study is one of the research designs that fell into the paradigm of ethnographic studies. The case study approach dealt with the process rather than the fixed data as in experimental studies. There were different levels of subjective involvement in this type of research. It was also contented that the differences between qualitative research and quantitative research are not dichotomous, but continual (Goetz and LeCompte, 1984, p. 50).

Emig's study was the first study in which the case study approach was used fully and systematically in the research of the writing process of high school students in the United States, though this approach had been widely used as a mode of inquiry within the physical, biological, and social sciences (Robinson, 1946; Piaget, 1955; Freud, 1956; Bettelheim, 1960). In her report, Emig cited an English writer, David Holbrook (1961), and an American teacher, Herbert Kohl (1967), as the only two persons who had used the case study approach in studying writing. She pointed out that her study "represented a unique effort to utilize case studies for eliciting data about how students behave as they write." She believed that the value of her study was in the fact that it steadily assumed that "persons, rather than mechanisms, compose." (Emig, 1971, p. 5) In applying the case study approach to the study of the writing process of high school students, the researcher hoped to illuminate

the psychological dimension of student writing (p. 2). According to Emig, the purpose of her study was "to attempt to describe how student writers usually or typically behave as they write with minimal direct intervention by the investigator." (p. 21)

Emig identified a group of eight sixteen- and seventeen-year-old secondary school students from different types of schools in the greater metropolitan area of Chicago who volunteered as subjects for her study. Of the five girls and three boys, six were white; one black and one Chinese-American. All the eight students were considered to have average or above-average intelligence, partly based on their school records and partly based on College Entrance Board Examination scores. Furthermore, some the subjects were recommended by their teachers as "good" writers and some subjects in the group were selected as NCTE Achievement Awards winners for their schools.

Each subject met with the investigator four times during the study. The first session consisted of twenty minutes of informal conversation and a period of free writing with the composing aloud technique by the subject. At the end of the first session, a stimulus of unspecified assignment was given for the second session. At the second session the subject was also required to compose aloud and tell all the prewriting and planning done between the sessions. At the end of the second session, the subject was

asked to recall all the writings he/she had done in the past by consulting parents and former teachers and to bring whatever writing samples he/she had. At the third session, the subject was requested to give a writing autobiography that included related reading experiences and descriptions and evaluations of the teaching of writing he/she had experienced. At the end of the third session, the subject was asked to bring a piece of imaginative writing to the fourth session with all the plans and drafts they did between the sessions. At the fourth session, the subject was asked to recount the process of their writing the imaginative piece. All the sessions had about one week between them. The investigator used a tape recorder to record the composing aloud of the subject and took notes of the subject's actions during the sessions.

The thirty-two sessions of working with the subjects accumulated a large amount of data. By dissecting the data gleaned from these sessions and focusing on the dimensions of the composing process, Emig produced a list of the dimensions of the writing process (Table 10).

In her discussion of the data and the dimensions of the student composing process, Emig used the schema of registers devised by the British linguists Halliday, McIntosh, and Strevens because of the applicability of their category-system to her inquiry (p. 36). In her effort to categorize the writing process, Emig concluded:

Table 10

DIMENSIONS OF THE COMPOSING PROSESS
AMONG TWELFTH-GRADE WRITERS

1. Context of Composing
Community, Family, School
2. Nature of Stimulus
Register:
Field of Discourse--encounter with natural environment; encounter with induced environment or artifacts; human relationships; self.

Mode of Discourse--expressive-reflexive; expressive-extensive.
Tenor of Discourse

Self-Encountered Stimulus

Other-Initiated Stimulus:
Assignment by Teacher--external features (student's relation to teacher; relation to peers in classroom; relation to general curriculum and to syllabus in English; relation to other work in composition); internal features or specification of assignment (registers, linguistic formulation, length, purpose, audience, deadline, amenities, treatment of written outcome, other).
Reception of Assignment by Student--nature of task, comprehension of task, ability to enact task, motivation to enact task.
3. Prewriting
Self-Sponsored Writing:

Length of period
Nature of Musings and Elements Contemplated--field of discourse; mode of written discourse; tenor or formulating of discourse.

Interveners and Interventions--self, adults (parent, teacher, other), peers (sibling, classmate, friend); type of intervention (verbal, nonverbal), time of intervention,

Table 10 (continued)

DIMENSIONS OF THE COMPOSING PROSESS
AMONG TWELFTH-GRADE WRITERS

reason for intervention (inferred), effect
of intervention on writing, if any.

Teacher-Initiated (or School-Sponsored)
Writing:

(Same categories as above)

4. Planning

Self-Sponsoring Writing:

Initial Planning--length of planning; mode of
planning (oral; written; jottings,
informal list of words/phrases, topic
outline, sentence outline); scope;
interveners and interventions.

Later Planning--length of planning; mode;

scope; time of occurrence; reason;
interveners and interventions.

Teacher-Initiated Writing:

(Same categories as above)

5. Starting

Self-Sponsored Writing:

Seeming Ease of Difficulty of Decision
Element Treated First Discursively--seeming
reason for initial selection of that
element; eventual placement in completed
piece.

Context and conditions under Which Writing
Began

Interveners and Interventions

Teacher-Initiated Writing:

(Same categories as above)

6. Composing Aloud: A Characterization

Selecting and Ordering Components:

Anticipation/Abeyance--what components
projected; when first noted orally; when
used in written piece.

Kinds of Transformational Operations--
addition (right-branching, left-
branching); deletion; reordering or

Table 10 (continued)

DIMENSIONS OF THE COMPOSING PROSESS
AMONG TWELFTH-GRADE WRITERS

substitution; embedding.

Style--preferred transformations, if any;
"program" of style behind preferred
transformations (source: self, teacher,
parent, established writer, peer); (effect
on handling of other components--lexical,
rhetorical, imagaic).

Other Observed Behaviors:

Silence--physical writing; silent reading;
"unfilled" pauses.

Vocalized Hesitation Phenomena--filler sounds
(selected phonemes; morphemes of
semantically-low content; phrases and
clauses of semantically-low content);
critical comments (lexis; syntax;
rhetoric): expressions of feelings and
attitudes (statements, expressions of
emotion--pleasure/pain) toward self as
writer to reader; digressions (ego-
enhancing; discourse-related).

Tempo of Composing:

Combinations of composing and Hesitational
Behaviors

Relevance of Certain Theoretical Statements
Concerning Spontaneous Speech

7. Reformulation

Type of Task:

Correcting; Revising; Rewriting

Transforming Operations:

Addition--kind of element; stated or inferred
reason for addition.

Deletion--kind of element; stated or inferred
reason for deletion.

Reordering or Substitution--kind of element;
stated or inferred reason.

Embedding--kind of element; stated or
inferred reason.

Table 10 (continued)

DIMENSIONS OF THE COMPOSING PROSESS
AMONG TWELFTH-GRADE WRITERS

8. Stopping
Formulation:
 Seeming Ease or Difficulty of decision
 Element Treated Last--seeming reason for
 treating last; placement of that element
 in piece.
 Context and Conditions under Which Writing
 Stopped

 Interveners and Interventions
 Seeming effect of Parameters and Variables--
 established by others; set by self.
Reformulation:
 (Same categories as above)
 9. Contemplation of Product
 Length of Contemplation
 Unit Contemplated
 Effect of Product upon Self
 Anticipated Effect upon Reader
 10. Seeming Teacher Influence on Piece
 Elements of Product Affected:
 Registers--field of discourse; mode of
 written discourse; tenor of discourse.
 Formulation of Title or Topic;
 Length; Purpose; Audience;
 Deadline; Amenities; Treatment of Written
 Outcome;
 Other. (Emig, 1971, pp. 34-35)
-

Source: Emig, 1971, pp. 34-35.

This chapter represents a theoretical sketch of one of the most complex processes man engages in. Although it is roughly taxonomic, it does not of course purport to be exhaustive. Nonetheless, almost every sentence contains or implies hypotheses upon which one could spend a lifetime in empirical research. Perhaps investigators other than the writer will find here materials for provocative questions and generative hypotheses about the composing process, particularly of students (p. 44).

Another case study included in this study was conducted by Swanson-Owens (RL) to examine the natural sources of resistance among high school teachers against the implementing of writing across the curriculum. Swanson-Owens had as her subjects two high school teachers, one English teacher and one home economics teacher, who volunteered with her project. The high school where this study was conducted was in an urban middle class community. Her approach in this case study was characterized by the researcher-teacher cooperation in an effort to narrow the gap between theory and practice, which frequently contributed to the limited success in curriculum reforms (Olson, 1980; Fulwiler, 1984).

The researcher met with the teachers for a forty-minute discussion every week for five months. The research also made 18 visits to each teacher's classroom during this period. All the interviews and observations were tape recorded, resulting in about 200 pages of text. The researcher also collected the following materials for this case study:

1. A writing assignment given at the beginning, middle and end of the school year.
 2. All writing-to-learn tasks given as a result of author's collaborative work with the teacher.
 3. Any assignment that seemed "interesting" for one reason or another--these included "unusual" assignments given by the teacher, and those deemed particularly successful or problematic by the teacher.
 4. All "writing" (defined as 3 or more connected sentences) done by two case study students in all of their other classes.
 5. Various materials used by teachers in planning their lessons.
 6. The English department curriculum guidelines.
 7. The teacher's comments on student papers.
 8. Worksheets used by students on the days the researcher observed the classes.
 9. Any written correspondence between the teacher and the researcher (e.g. for a while the teachers were keeping a journal).
- (Swanson-Owens, 1986, p. 74).

In the analysis of the data, Swanson-Owens developed a Model of Curricular System of Meaning based on Elbaz's (1981) analytical categories as a frame. The model described in RL was a matrix of four factors: material/activity, teacher, knowledge, and students. The four factors devised in this study matched Eisner's notions of planned events, intention, educational consequence, and students (Table 11).

The analysis of the data by means of this model revealed that the teachers had different views and concerns

about the factors identified in this matrix. Their responses to the new curricular elements introduced were based on their individual views and concerns in regard to this matrix. What was called "resistance" by an outsider researcher in implementing curricular renovation was found to be derived from the different views and concerns of the individual teachers. RL identified two aspects for careful examination in future studies--locus of attention (which specified a point of vulnerability in the system), and conditions of instruction (which affected how instruction took place) (p. 94). RL also offered this model of analysis for studies in other areas of interest and claimed that this analytical model could promote a dialogue between researchers and teachers about some fairly fundamental ideas which influence curricular decisions and instructional practices (p. 95).

Assessment, Measurement, and Evaluation Studies

Of the twenty research reports selected for this study, Mellon's study (1975), Steward's study (1978), Grobe's study (1981), McCulley's study (1985), Chesky and Hiebert's study (1987), and McCann's study (1989) belonged to this group. Before we begin the analysis of these research reports, it is necessary to discuss briefly the differences and similarities among these three types of research. An

Table 11

MODEL OF A CURRICULUM SYSTEM OF MEANING

	Knowledge	Materials/ Activities	Teacher	Student
Knowledge	What is knowledge? What do the teachers want the students to learn?			
Materials/ Activities	How do the materials/activities contribute to learning?	What are the materials? What are the activities?		
Teacher	What does the teacher know about the subject matter?	What does the teacher do with materials (e.g. How does he/she set up, carry out, evaluate activities)?	What is the role of a teacher? *self-image *professional image	
Students	What do students know in general? about the subject matter? How do they respond to information?	How do students interact with the materials? What can they handle? How do they cope?	What is the relationship between the teacher and the students? *social *academic	Who are the students? *attitudes *qualities *characteristics *interests

Source: Swanson-Owens, 1986, p. 77.

assessment is defined as the systematic process of collecting and analyzing data in order to indicate the extent to which individual items of knowledge are known by persons in general (National Assessment of Educational Progress, in Mellon, 1975, p. 10). Measurement is defined as the systematic process of collecting and analyzing data in order to determine the norm and distribution of a given population. Evaluation is defined as the systematic process of collecting and analyzing data in order to make decisions on alternatives (Gay, 1987, pp. 7-8). Based on these definitions, Mellon's study was an assessment study. The other studies were measurement studies. No evaluation study was included here.

Mellon's study was known as the National Assessment of Educational Progress (NAEP), a federally funded survey of the educational attainments of representative samples of American young people and adults at four age levels. The survey was conducted in ten subject areas: reading, writing, literature, science, mathematics, citizenship, music, social studies, art, and career and occupational development. In each area, four age groups were assessed: nine, thirteen, seventeen, and adults of 26-35, matching the fourth grade, the eighth grade, the twelfth grade, and ten years beyond high school. There were 2,000 - 2,500 respondents in each group for this assessment, proportionately representing the entire nation at the age level in question. These

respondents were compiled also in terms of the following subgroups: geographical region (national quadrants: southeast, northeast, central, and west), sex, color (nonblack and black), parents' highest educational level (four categories: no high school, some high school, high school graduate, beyond high school), size and type of community (seven categories: extreme inner city, extreme rural, suburban, extremely affluent suburbs, small city, medium-sized city, the rest of big cities) (Mellon, 1975, pp. 8, 11, 12). This assessment was supposed to be conducted once every four years. Mellon's study discussed only three areas: writing, reading, and literature. For this study, only the writing part in the National Assessment of Educational Progress will be discussed.

In the writing assessment, four objectives were determined by a panel of seven educators based on an initial proposal of five categories offered by Educational Testing Service (ETS). Following were the four objectives in the final form:

1. write to communicate adequately in a social situation)letters, directions, formal notes, addressing envelopes, invitations);
2. write to communicate adequately in a business or vocational situation (information and application forms, mail order letters, business invitations, formal letters);
3. write to communicate adequately in a scholastic situation (notes and announcements; narrative, descriptive, and expository essays);

4. appreciate the value of writing (recognize the value of writing; write as a normal course of behavior; receive satisfaction from writing).

(pp. 14-15)

Based on this group of writing objective, ETS personnel under NAEP supervision formulated a body of exercises measuring the extent to which each objective had been attained by students at the several age levels. Three kinds of exercises were actually used:

1. acceptable/unacceptable exercises: short writing tasks scored "acceptable" or "unacceptable" according to whether or not the writer included certain required items of information; used with objectives one and two;
2. yes/no exercises: questions of beliefs or facts responded to by a "yes" or "no"; used with objective four;
3. essay exercises: topics requiring fully developed compositions, each rated by two trained readers on a scale of 0 (blank page) to 8 (best) on the basis of overall quality; used with objective three.

(p. 15)

Altogether, there were ten acceptable/unacceptable exercises, nineteen yes/no questions, and three essay exercises for nine-year-olds and thirteen-year-olds and two essay exercises for seventeen-year-olds. All exercises were administered by means of written directions accompanied by the playing of a tape-recorded reading of their text. In Appendix C, examples of the exercises used in Mellon's study was provided.

In rating the essays, two trained readers rated by

means of holistic scoring each essay for its overall quality using a 0 (blank page) through 8 (highest quality) scale. By adding the two scores together, a range of 0 - 16 was produced. In studying the writing mechanics, three kinds of procedures were employed in characterizing the mechanics of these essays:

1. computer counts of 24 quantifiable features of texts, such as average word and sentence lengths, numbers of each kind of punctuation mark, misspellings of certain words the computer had been programmed to look for, and so on;
2. generalized prose characterizations of specific mechanical strengths and weaknesses of writing typical of the three quality levels at each age, prepared by panels of recognized experts;
3. error counts of each of eight familiar types of errors, performed by teams of experienced English teachers. Two teachers working independently counted a given error type in each paper, and their totals were averaged.

(p. 30)

In the actual rating, the acceptable/unacceptable scoring approach was soon replaced by the holistic scoring approach due to the imprecise wording of the essay topics and the rhetorical diversity of the writing elicited. In the process of this assessment, no norm or absolute standards were invoked. In short, the Mellon study presented the actual picture of the writing ability of the four age groups of the nation based on the first national sampling of nearly 100,000 Americans (Mellon, 1975, p. 30).

Stewart's study (REC), a research project conducted by a Canadian professional, examined the syntactic maturity from high school to university based on the findings of two sets of studies by American researchers, one by Hunt (1965 and 1970), who published his influential report on differences in grammatical structures written at three grade levels analyzed by the transformational methods and one by O'Donnel, Griffin, and Norris (1967), who also used the transformational methods in their study of the syntax of kindergarten and elementary school children. Hunt's studies focused on the free writing of the school children of the 4th, 8th, and 12th grade and the writing of "skilled" adults and proposed the "minimal terminable unit" or "T-unit" as the most significant index of syntactic maturity for school groups. O'Donnel, Griffin, and Norris focused on the sampled speech of children in grades 1, 2, 3, 5, and 7 and their free writing in grade 3, 5, and 7 in the light of Hunt's T-unit approach and found a progressive increase in the average length of this measure through every grade (cited in Stewart, 1978). The measurement instruments developed by Hunt were used widely in composition studies both in and out of the United States. Stewart's study was one of the studies in the latter category, in which Stewart decided to use Hunt's T-unit and other measurements to study the syntactic maturity of students from high school to university. He used the same instrument "Aluminum"

developed by Hunt (1970) on 126 students of grades 10, 11, and 12 in Fredericton High School, Fredericton, New Brunswick, Canada and 176 Faculty of Education students at the University of New Brunswick in the same city. Stewart used the already established groups in the high school, one regular and one upper-level English class from each of the three grades that participated. The number of the students were reduced to 21 for each group through randomized selection. All the university students who volunteered in this study were included, resulting in groups of 40, 50, 50, and 35 for first-, second-, third-, and fifth-sixth year education sections. The university students were all supposed to be from regular and upper-level and college preparatory classes. Therefore, some degree of continuity was assumed between the high school population and university population under investigation (Stewart, 1978, p. 40). The results of this measurement study were in line with Hunt's and O'Donnel, Griffin, and Norris's studies. A progressive increase in syntactic maturity was found that closely corresponded to the results in Hunt's and O'Donnel, Griffin, and Norris's studies respectively (Tables 12, 13).

The advantages and disadvantages of using the instrument of "Aluminum" estimated by Hunt in 1970 were supported in this study as follows:

1. It is easily administered and scored, and the ratings obtained can be readily compared with

Table 12

MEAN SCORES ON THREE FACTORS OF SYNTACTIC MATURITY
O'DONNELL, 1967, AND HUNT, 1965 AND 1970

	G3	G4	G5	G6	G7	G8	G10	G12	AA	SA
O'Donnell (1967)										
W/C	6.50		7.40		7.70					
C/T	1.18		1.27		1.30					
W/T	7.67		9.34		9.77					
Hunt (1965)										
W/C	6.60					8.10		8.60		11.50
C/T	1.29					1.42		1.68		1.74
W/T	8.51					11.34		14.40		20.30
Hunt (1970)										
W/C	5.19		5.76			6.79	7.35	7.85	8.40	9.95
C/T	1.04		1.18			1.43	1.41	1.44	1.47	1.51
W/T	5.42		6.84			9.84	10.44	11.30	11.85	14.78
W/C words per clause						G grade				
C/T clauses per T-unit						AA average adult				
W/T words per T-unit						SA skilled adult				

Source: Stewart, adapted from Hunt and O'Donnell.

Table 13

MEAN SCORES ON THREE FACTORS OF SYNTACTIC MATURITY
GRADE TEN THROUGH SIXTH YEAR UNIVERSITY

	G10	G11	G12	U1	U2	U3	U5-6
Stewart (1978)							
W/C	7.55	8.22	8.56	8.25	7.92	8.16	8.81
C/T	1.44	1.50	1.50	1.48	1.57	1.55	1.60
W/T	10.83	12.32	12.87	12.25	12.36	12.61	13.85
W/C words per clause				G grade			
C/T clause per T-unit				U university			
W/T words per T-unit							

Source: Adapted from Stewart, p. 41.

results of earlier studies which provide tentative norms.

2. It gives a significant difference at a two-year interval in high school and between high school and senior university.
3. Its reliability has not been established through a second administration to the same groups or through a comparison of "Aluminum" scores with free writing scores of the same students. However, Hunt (1970) did find some correlation between "Aluminum" scores and scores on another instrument of a similar sort.
(Stewart, 1978, p. 46)

Grobe's study (RHGC) reexamined the indexes of syntactic maturity in student writing, but for a different purpose. The researcher tried to correlate the indexes with the holistic quality rating by the teachers. Using the database for 1977 New Brunswick (Canada) Writing Assessment Program (WAP), the researcher in a previous study found that, as predictors of quality, composition length and freedom from spelling errors accounted for about 25% of the variance in teacher quality ratings. In grade 5, syntactic maturity variables explained about 10% of the accountable variance, while at grades 8 and 11, they contributed virtually nothing to the prediction of quality rating. This study, therefore, attempted to examine the relationship between specific vocabulary characteristics and teacher quality ratings (Grobe, 1981, p. 75).

The 1979 WAC was participated by approximately 7,500 5th graders, 8,000 8th graders, and 6,700 11th graders. While the 1977 WAC presented an expository writing

instrument, the 1979 WAC presented a narrative writing instrument. The writing task was presented in the following ways:

For the 5th Grade:

Imagine you found an envelope with \$100 in it on the road one day. You tried hard, but could not find the owner. Your parents wanted you to save the money, but agreed that you could spend it if you wished. Write a story of what happened, starting with the finding of the money!

For the 8th Grade:

Students were presented with a cartoon showing a young boy sitting at a restaurant table with an adult. The boy is saying, "Do I have to get the child's plate? I'm SEVEN!" Following the

cartoon, the writing assessment item read: Select one time when you, or someone you know, was in a similar situation. Make this into a story.

for the 11th Grade:

The same cartoon was presented, but the test item was revised to read:

The young man in the cartoon is annoyed with his mother because she fails to recognize that he is growing up. However, this is an experience which is not limited to young children.

Write a story based on this kind of conflict and drawn from your own observation and experience.

Readers randomly selected 950 papers from grade 5, over 1,000 papers from grade 8, and over 750 papers from grade 11. Then the samples were holistically scored by two independent raters on a 1 to 4 point scale (4 being the highest award). If the two ratings differed by more than one single point, the paper was read by a third person,

usually the chief reader, to resolve the difference.

As in the 1977 study, a second stage sample was randomly selected: 153 from the 5th grade papers, 151 from the 8th grade papers, and 133 from the 11th grade papers. The mean and standard deviation of the second stage sample were reported to be the same as those of the initial sample (p. 77). To analyze the second stage sample, this researcher identified the following independent variables:

1. W--the total number of words written
 2. S--the number of spelling errors per 100 words
 3. PARA--paragraphing present (1=yes, 0=no)
 4. STANCE--third person (1=third person, 0=not third person)
 5. W/T--words per T-unit
 6. W/C--words per clause
 7. C/T--clauses per T-unit
 8. F/T--fragments per T-unit
 9. R/T--run-ons per T-unit
 10. U/T--faulty pronoun references per T-unit
 11. AAC--adverb-adjective confusion per T-unit
 12. CC--case errors per clause
 13. VERB--form errors per clause
 14. NA--non agreement per T-unit
- (Grobe, 1981, p. 77)

After a stepwise regression of holistic score was applied on the 14 variables, the results showed that for the 5th graders, composition length and spelling were the two "best" indicators of quality rating, accounting for 45.4% of the variance for 5th grade, 30.4% of the variance for 8th grade, and 25.4% of the variance for 11th grade. However, still a large percentage of the variance was left unaccounted for. For the purpose of increasing the accounted for proportion of holistic score variance, more

variables related to vocabulary characteristics were identified and a third stage sample was randomly selected, approximately 50 from each group. The additional vocabulary variables identified were as follows:

15. FLESCH--Flesch's Reading Ease score for 50 per cent comprehension. It is based on T-unit length and syllable size.
16. DIVE--an index of vocabulary diversity, the reciprocal of YULE'S K.
17. TYPE--the number of different words which appeared in the composition.
18. RANK--the percentage of words which were from a list of the 500 most frequently used words.
19. YULES--a measure of vocabulary diversity which is relatively independent of sample size and content.
20. TTR--the type-token ratio
21. POWERS--a readability measure very similar to Flesch's which yields a grade equivalent score interpreted as a Dale-Chall readability equation.
22. REPEAT--the vocabulary repeat rate
23. WDSIZE--the total number of letters in the essay divided by the number of words.
24. VOCDIV--a measure of vocabulary diversity.

$$\text{Vocdiv.} = \frac{\text{S.D. of Repeat Rate}}{\sqrt{\text{TYPES}}}$$

The third stage sample was found to be parallel to the second stage sample in terms of mean and standard deviation (p. 80). The addition of the vocabulary variables increased the accounted for percentage of holistic score variance from 59.8 to 82.4 (p. 80). Most of the vocabulary characteristics variables were found to be more important in accounting for the holistic quality scores variance than the syntactic maturity and mechanics variables (Tables 14, 15, 16).

Table 14

PREDICTORS OF GRADE FIVE HOLISTIC QUALITY SCORE, n=45

Step	Variable	Cumulative	Per Cent	R
		Per Cent of Accountable Variance	Unique Contribution to Accountable Variance	
1.	S	51.9	51.90	.519
2.	TYPE	55.4	3.50	.554
3.	PARA	60.5	5.00	.778
4.	RANK	62.5	1.90	.790
5.	REPEAT	64.4	1.90	.802
6.	VOCDIV	66.4	1.90	.815
7.	W	67.3	0.90	.820
8.	TTR	68.9	1.50	.830
9.	FLESC	70.6	1.70	.840
10.	WDSIZE	75.2	4.60	.867
11.	AAC	76.3	1.00	.873
12.	POWERS	77.1	0.80	.878
13.	NA	78.2	1.00	.884
14.	CT	79.2	0.90	.889
15.	YULES	79.6	0.40	.892
16.	DIVE	80.0	0.30	.896
17.	W/T	80.3	0.30	.902
18.	W/C	81.4	1.10	.905
19.	PRT	81.9	0.50	.906
20.	F/T	82.2	0.50	.907
21.	V	82.3	0.10	.907
22.	R/T	82.3	0.00	.907
23.	CC	82.4	0.00	.907
24.	STANCE	82.4	0.00	.907

Maximum adjusted R^2 obtained.

Table 15

PREDICTORS OF GRADE EIGHT HOLISTIC QUALITY SCORE, n=50

Step	Variable	Cumulative	Per Cent	R
		Per Cent of Accountable Variance	Unique Contribution to Accountable Variance	
1.	S	39.8	39.80	.631
2.	POWERS	50.6	10.70	.711
3.	W	54.2	3.50	.736
4.	R/T	56.0	1.80	.748
5.	YULES	57.5	1.40	.758
6.	VOCDIV	58.7	1.10	.766
7.	PRT	62.6	3.90	.791
8.	FT	63.8	1.10	.799
9.	DIVE	64.9	1.10	.805
10.	NA	65.3	0.40	.808
11.	V	65.5	0.20	.809
12.	RANK	65.8	0.20	.811
13.	STANCE	65.9	0.10	.812
14.	AAC	66.2	0.10	.813
15.	W/T	66.3	0.20	.814
16.	FLESCHE	66.4	0.10	.815
17.	TYPES	66.4	0.00	.815
18.	REPEAT	66.7	0.30	.817
19.	TTR	68.7	*2.00	.829
20.	CC	69.0	0.20	.830
21.	PARA	69.1	0.101	.831
22.	WD	69.2	0.00	.832
23.	C/T	69.2	0.00	.832

Maximum adjusted R^2 obtained.

* Original is 0.20, a typing error.

Table 16

PREDICTORS OF GRADE ELEVEN HOLISTIC QUALITY SCORE, n=49

Step Variable	Cumulative		Per Cent Unique Contribution to R
	Per Cent of Accountable Variance	Accountable Variance	
1. TYPES	29.3	29.30	.541
2. S	38.0	8.70	.617
3. DIVE	46.4	8.40	.681
4. VOCDIV	49.5	3.10	.704
5. FLESCHE	53.5	3.90	.731
6. STANCE	55.6	2.10	.746
7. CC	57.6	1.90	.758
8. PRT	58.7	1.10	.766
9. RANK	59.9	1.10	.773
10. W	61.0	1.10	.781
11. R/T	62.0	1.00	.787
12. YULES	62.7	0.60	.791
13. WDSIZE	62.6	0.90	.797
14. POWER	65.8	2.10	.811
15. AAC	66.5	0.69	*.815
16. W/T	67.2	0.73	.820
17. F/T	67.4	0.16	.821
18. W/C	67.5	0.14	.822
19. C/T	68.5	0.96	.827
20. REPEAT	68.8	0.28	.829
21. TTY	70.3	1.56	.838
22. V	71.7	1.35	.847
23. PARA	71.7	0.03	.847

Maximum adjusted R^2 obtained.

* Original is .811, a typing error.

Source: Grobe, 1981, pp. 82-84.

In conclusion, Grobe stated that although vocabulary diversity was found to be the "best" predictor of holistic scoring for narrative writing, a successful narrative was still composed of complex sets of multivariate effects which act together in a presently unrecognized fashion and it was a mistake to concentrate on improving students' vocabulary in order to improve their writing (p. 85)

McCulley's study (RK, 1985) investigated the relationships among features of textual cohesion, as identified by Halliday and Hasan (1976), and primary-trait assessment of writing quality and coherence, with manuscript length held statistically constant. The sample selected for this study was from the 1978-1979 National Assessment of Educational Progress (NAEP) writing section.

McCulley based his framework of study on the works of Bain (1867), Kineavy (1971), Halliday and Hasan (1976), De Beaugrande (1980), De Beaugrande and Dressler (1981), and the results from the writing section of the first (1969-1970), the second (1974-1975), and the third (1978-1979) NAEP studies. RK made the distinction between coherence and cohesion based on studies by De Beaugrande and Dressler (1981) and Witte and Faigley (1981) where coherence was defined as "not a mere feature of texts, but rather the outcome of cognitive processes among text users" (De Beaugrande and Dressler, 1981, in McCulley, 1985, p. 270), and cohesion was defined as "a product only of the factors

operating within the text that bind it together" (McCulley, p. 270).

For textual cohesion, RK referred to Halliday and Hasan's study (1976) for a cohesion taxonomy (Table 17), which was not exhaustive according to Markels (1983) and Stotsky (1983b), but comprehensive enough for this study (McCulley, 1985, p. 270). On the other hand, RK used the NAEP primary-trait coherence measures in Table 17 in its effort to determine the relationships between the two types of measures. Table 18 shows the NAEP scoring guide.

McCulley used a random sample drawn from the NAEP parent sample which was drawn based on a stratified randomization on sex and geographical locations of the 17-year-old population of the United States attending school in the 1978-1979 academic year, excluding non-English-speaking persons, those identified as nonreaders, persons physically or mentally unable to respond, and persons in institutions or attending schools established for the physically or mentally handicapped (NAEP, in McCulley, 1985, p. 271). Of the original 493 papers, twenty were randomly selected from each of the six groups that were rated 1-2, 3, 4, 5, 6, 7-8 respectively on a 1-8 scale, making a total of 120 papers in this subgroup. All the papers were on the same topic as follows:

Table 17

HALLIDAY AND HASAN COHESION CATEGORIES

Category	Example
1. Reference: Personal Pronoun	"he"
2. Reference: Demonstratives	"the"
3. Reference: Comparatives	"more"
4. Substitution and Ellipsis: Noun	"one"
5. Substitution and Ellipsis: Verb	"do"
6. Substitution and Ellipsis: Clause	"so"
7. Conjunction: Adding	"and"
8. Conjunction: Contrasting	"yet"
9. Conjunction: Cause	"because"
10. Conjunction: Time and Summation	"next"
11. Conjunction: Continuing	"well"
12. Lexical: Same	"car & car"
13. Lexical: Synonym/Hyponym/Collocation	"car & automobile"
14. Lexical: Higher Category	"car & vehicle"
15. Lexical: General nouns	"car & thing"
16. Total number of cohesive ties	
17. Total number of different types of cohesive ties	

Source: Halliday & Hasan (1976), in McCulley, 1985, p. 282.

Table 18

NAEP SCORING GUIDE: PRIMARY TRAIT COHERENCE (COHESION)

Scoring Guide Categories:

- 1 = Litter or no evidence of cohesion. Basically, clauses and sentences are not connected beyond pairings.
- 2 = Attempts at cohesion. There is evidence of gathering details but little or no evidence that these details are meaningfully ordered. In other words, very little seems lost if the details were rearranged.
- 3 = Cohesion. Details are both gathered and ordered. Cohesion is achieved in the ways illustrated briefly in the definition above. Cohesion does not necessarily lead to coherence, to the successful binding of parts so that the sense of the whole discourse is greater than the sense of its parts. In pieces of writing that are cohesive rather than coherent, there are large sections of details which cohere but these sections stand apart as sections.
- 4 = Coherence. While there may be a sense of sections within the piece of writing, the sheer number and variety of cohesion strategies bind the details and sections into a wholeness. This sense of wholeness can be achieved by a saturation of syntactic repetition throughout the piece and/or by closure which retrospectively orders the entire piece and/or by general statements which organize the whole piece.

Source: NAEP (1980), in McCulley, 1985, p. 280.

Recreation Center

Some high school students have proposed converting an old house into a recreation center where young people might drop in evenings for talk and relaxation. Some local residents opposed the plan on the grounds that the center would depress values in the neighborhood and attract undesirable types. A public hearing has been called. Write a brief speech that you would make supporting or opposing the plan. Remember to talk only ONE point of view. Organize your arguments carefully and be as convincing as possible. Space is provided below and on the next three pages.
(NAEP, 1980, p.25, in McCuilley, 1985, p. 272)

The raters for the parent sample were four full-time faculty members from the Utah State University English department trained for the rating with an interrater reliability of .96. The subsample of 120 papers was rated by the research and a graduate assistant using the Halliday and Hasan categories after training with an interrater reliability of .91 (McCulley, 1985, p. 273).

With the length of the papers (number of the T-units) held constant, partial correlation coefficients were calculated for all relationships. Then multiple regression was used to examine the maximum contribution of cohesion, as defined by Halliday and Hasan, to the primary-trait measures and coherence measures of the persuasive writing for authoritative audience. The results were as follows (Table 19):

Chesky and Hiebert's report (RM2) of 1987 was based on Chesky's doctoral dissertation, a study of the effects of low- and high-prior knowledge and peer and teacher audiences

Table 19

BEST POSSIBLE SUBSET FROM REGRESSION ON THE PRIMARY-TRAIT

Variable	Coefficient	Standard Error	T-Ratio	Significance	Unique Contribution to R ²
T-units	.14	.03	5.03	$p < .001$	.11
Conjunction: Contrasting	3.80	1.58	2.41	$p < .018$	.02
Lexical Repetition	.54	.20	2.92	$p < .004$	.03
Lexical Synonym, Hyponym, & Collocation	1.34	.25	7.31	$p < .001$	.17

Source: McCulley, p. 277.

on high school students' writing. Research in the effect of prior knowledge on writing began in the past decade with Hilgers's study (1982) and Langer's study (1984). Hilger's study began by examining the relationship between the Scholastic Aptitude Test scores of college students and the quality of their essays (with a mean of .29 across the nine measures adopted). But when Hilgers controlled for familiarity of content matter, the correlation became negligible (.04). Hilgers concluded that topic-specific prior knowledge might exert a stronger influence on a student's writing than any particular training program or any research variable, such as general knowledge, ethnicity, previous education, or gender (in Chesky & Hiebert, 1987, p. 304). Langer correlated a number of quantitative and qualitative measures to students' prior knowledge in her

study and found significant correlations between overall essay quality and three measures of topic-specific prior knowledge, ranging from .20 to .30 (p. 305).

The subjects in RM2 were eighty students in four 11th grade classes of a high school in a small midwestern town. Two classes were practical English classes and two were academic English classes, respectively representing the lower half and the upper half of the student population. A 2 (Prior Knowledge: low, high) X 2 (Audience: teacher as examiner, peers) factorial design was used. Twenty students were randomly assigned to each of the four groups after they had taken a pre-test on the prior knowledge on the composition topics. The prior knowledge test was a word association test designed based on Langer's study (1984) and Newell's study (1984) in accordance with Meyer's scheme (1975). Two trained raters who were former teachers and current graduate students, read related articles independently and selected and rank ordered five concept words central to each of the themes. The interrater reliability was .75 (p. 306). A consensus approach was adopted to determine the first three most important concept words for each of the four themes for the study. A pilot study was conducted before this study.

The students were given a sample word association test before they took this pre-test, which was given in different but equalized forms. The pre-test was scored on three

dimensions of fluency, organization, and combination by two trained raters whose interrater reliability was .98 for fluency, .87 for organization, and .95 for combination.

Based on the results of this pre-test, two themes which showed the greatest discrepancy in scoring were selected for the writing study, one for low prior knowledge (Tobacco Price Supports) and one for high prior knowledge (The Problem with Teachers). To control for general language ability and gender, the boys and girls in each of the four classes were rank ordered separately. Then the upper half of each group and the lower half of each group were separately assigned to the four groups in the 2 X 2 factorial design by means of randomization. No language difference was found between the groups using scores on the language portion of the California Test of Basic Skills.

Three weeks later, the four group were given four different packages for the writing test based on the 2 X 2 design. Each group had 60 minutes for the writing. After the test, the students were administered an affect survey with a five-point Likert scale (1932) on their feelings and attitudes concerning the writing.

Two trained raters rated the four groups of writings on six variables: overall quality (holistic scoring), essay length, context-creating statements, cohesion, syntactic complexity (mean T-unit length), and error analysis, where the interrater reliability was .83 for holistic scoring, .99

for essay length, .79 for cohesion, .87 for context-creating statements, .99 for syntactic complexity, and .97 for error analysis (p. 307).

RM2 used Myers's procedures (1980) for the holistic scoring. The essay length was the number of words for each essay. Content-creating statements were defined as "the range of information to be covered in the composition without imparting specific details" (Bracewell, Scardamalia, & Bereiter, 1978). Cohesion used the NAEP procedures in Mullis and Mellon (1980). T-unit and error analysis was also based on the Mullis and Mellon's study (1980).

After a 2 X 2 analysis of variance and the Pearson product-moment correlation coefficients were computed, it was found that prior knowledge had effects on holistic scores, essay length, and context-creating statements (Table 20).

McCann's study (RO) of 1989 examined the ability of students of grades 6, 9, and 12 to write argumentative prose. According to McCann, recent reports continued with the dismaying information about students' inability to write persuasive or argumentative essays (NAEP, 1986, In Applebee, Langer, and Mullis, 1986). Moffet (1968a) believed it was too abstract for young people. Bereiter and Scardamalia (1982) argued that students used oral discourse schemata in writing and their appropriate written discourse schemata waited to be developed. Crowhurst's three studies (1981,

Table 20

CORRELATIONS OF PRIOR KNOWLEDGE AND WRITING MEASURES

Writing Measure	Fluency	Organization	Combination
Holistic	.30***	.37***	.44***
Essay Length	.41***	.51***	.48***
Context-creating statements	.11	.27**	.20*
Cohesion	-.02	.19*	.17
T-unit	.00	.08	.17
Errors	-.12	-.13	-.23*

Source: Chesky and Hiebert, 1987, p. 310.

1983a, 1983b) studied the differences between elementary students and high school students in their argumentative writing, using Halliday and Hasan's analysis of elements in text cohesion in the 1981 study (in McCann, 1989, p. 64). RO focused on the knowledge of arguments and ability to produce argument of the students of grades 6, 9, and 12. The subjects of RO were 33 six graders, 40 ninth graders, and 22 twelfth graders, making a total of ninety-five from a large suburban high school and its feeder school. In addition, 60 university professors and NCTE writing committee members were contacted with the identification of the argumentative passages, of whom 22 responded as adult input into the study for comparison. In designing the instruments, Toulman's (1958) model of argumentation was used. Variables in argumentation were defined as claims, data, warrants, propositions, opposition, and response to the opposition. A list of features of argument was

generated to assess subjects' recognition of the argument features as follows:

1. No explicit features of argument, narrative only
2. Claim (proposition) plus narrative
3. Set of claims (propositions)
4. Claim plus data
5. Proposition, data, warrant
6. Proposition, claim, data, warrant
7. Proposition, claim, data, warrant, opposing point-of-view.

(McCann, 1989, p. 66)

With these features, seven passages were used representing each of the seven types listed above for the subjects to identify: food, horizon, seat belt, smoking, asbestos, poor, and violence (pp. 73-74). Besides this instrument, a writing assignment was designed for the subjects of the three grades to compose an argumentative essay. The writing prompt was as follows:

Recently, members of the student council at Franklin Pierce Memorial High School held a meeting. They discussed the school's policy which forbids students from leaving the school grounds during the school day. The only students allowed to leave the building are those who have been excused to go home because of some emergency. Some student council members want to change the rule to allow students to leave school during the lunch periods. During the meeting one student pointed out the principal's reasons for the rule. She said that many homeowners who live near the school are afraid that their property will be damaged. In addition, students might be injured when they are away from school. Some students

insist that they are old enough to leave the building on their own. They claim that the things that the principal is worried about would never happen.

Imagine that you are a member of the student council. Write a letter to the school principal. Explain whether or not students should be allowed to leave the building during lunch periods. Be sure to consider all possible reasons in support of your position. Explain each reason fully.

(p. 66)

Two raters were trained to rate all the papers. The scoring guide used was based on Toulman's model of six argumentative traits: claims, data, warrant, proposition, recognition of opposition (qualification), and response to opposition (rebuttal) (p. 67). As it was believed that backing would not appear in the students' compositions, backing was not included in this study.

Chi-square tests and one-way ANOVA were run to determine the significance of the difference between the age groups. Results showed that all age groups had some knowledge about argumentation, but the difference between the age groups were apparent. In the compositions, all age groups could make claims and state propositions, but the use of data, warrants, qualifications, and rebuttals was weak across the board (p. 72).

Strengths, Weaknesses, and Applicability

A large number of influential works had been published on these descriptive research methodologies. Concerning survey methodology, works by Hansen, Hurwitz, and Madow, 1953; Hyman, 1955; Lazarsfeld and Rosenberg, 1955; Butterfield, 1960; Kuhn, 1962; Stouffer, 1962; Cochran, 1963; Kish, 1965; Oppenheim, 1966; Glock, 1967; Rosenberg, 1968; Wallace, 1971; and Babbie, 1973 already gave a fairly detailed picture of the enterprise.

For Chinese professionals, in their dealing with the survey methodology, the following guide to a better comprehension of survey methodology is very helpful (Babbie, 1973, pp. 364-365):

1. What general topic does the researcher wish to examine?
2. What was his motivation in designing and executing the study?
3. Is the researcher's main purpose one of exploration, description, explanation, or a combination of these?
4. What general population are the findings meant to represent?
5. What was the sampling frame used for purposes of selecting a sample to represent that population?
6. How was the sample actually selected?
7. How many respondents were initially selected in the sample and how many actually participated?

8. To what extent is sampling error likely to affect the results of the survey?
9. How were the data collected? Was the data collection method appropriate to the population and subject matter of the study?
10. When did the data collection take place and how long did it take?
11. How were specific variables measured in the analysis? How were questionnaire items worded and/or how were responses combined into composite measures of variables? Does a given response pattern--imagining the orientations of persons giving it--reflect our common sensible interpretation of the summary term associated with the response pattern (for example, "very rigorous")?
12. Have composite measures been validated in such a way as to further ensure their adequate representation of the variables under consideration?
13. Are the methods of analysis used in the report appropriate both to the subject matter and the form of the data collected?
14. How strong are the associations discovered among variables? What do associations of that strength mean in understanding the real world?
15. Has the researcher adequately presented the logic of associations discovered empirically? Has he presented plausible reasons for those associations?
16. Has the researcher adequately tested for alternative explanations? Has he tested the possible spuriousness of the relationships?
17. Do the empirical relationships suggest further analyses that the researcher has neglected?
18. Could an independent reader replicate the survey on the basis of information presented in the report?

(Babbie, 1973, pp. 364-365)

Observational study had its root in studies by Benedict, 1934; Hall, 1959 and 1966; Glaser and Strauss, 1967; Manis and Meltzer, 1967; Blumer, 1969; Spradley and McCurdy, 1971 and Spradley, 1980; Lofland, 1976; Agnew and Pyke, 1978; and Bogden and Biklen, 1982.

Observation was common in Chinese schools. However, techniques in observational studies need improvement. The usual practice was mostly functional and occasional, where the objectives and instruments have not been carefully defined. Effort should be made to develop insight of how observation study can best be carried out and for what purpose.

Measurement studies are beginning to be carried out as a means to promote competition in Chinese schools. The test scores are usually taken very seriously, sometimes to the degree that the scores are decisive for the future life-long career of those who participate in the measurements. The notion of "absolute equality in front of the score" has been, and still is, a deep-rooted concept among the general public in dealing with the scores. Therefore, in implementing measurement studies in China, additional care has to be taken to put the objective of the test in the right place.

Statistics expertise is an extreme weakness among faculty of English departments. While faculty members of the education departments have training in statistics, few

of them have proficient working knowledge of English to undertake research in English composition. It is the author's expectation to develop a program for faculty and graduate students of English departments in China to acquire a working knowledge of educational statistics either through in-service training sessions or regular graduate programs.

Research methodology is the most important aspect of this anticipated program. Past research in the English curriculum in China has been mostly related to discussions about literature, issues in grammar, renovations of instructional techniques in classrooms, which usually showed a certain degree of subjective judgement. Reference has been very limited in their reports, partly because research was still scattered and unorganized, information system still in need of improvement, partly because few professionals have been systematically trained in research methodology.

In conclusion, this chapter discussed the selected reports which belonged to the category of descriptive studies, including survey studies, observation studies, assessment studies, and measurement studies. Methodological concerns and the unique situations in China were also discussed in the hope that a program could be developed to help Chinese professionals join the research enterprise to make their contributions to the academic world in the field of English composition research.

CHAPTER V

META-ANALYSIS

An Overview of Meta-Analysis

Criticism about empirical studies in social and behavioral sciences was phenomenal. Not only was the research effort along this line termed "sadly dilapidated," and facing a "crisis" (Glass, 1977; Jackson, 1980; Rosenthal, 1984), it was also accused of leading to an "enormous waste of scholarly effort in the social sciences" (Kulik, 1983). When science was looked upon as the accumulation and refinement of information and knowledge (Hunter, Schmidt, and Jackson, 1982), then the establishment of a set of guidelines for reliable and valid reviews, integrations, and synthesis of studies examining similar research questions became critical (Cooper, H. M., 1982).

The statistical methods, proposed by Glass, McGaw, and Smith (1981) and known as meta-analysis, regarded the findings of each individual study as one data point. Therefore, the results of a large number of studies addressing similar research questions were regarded as a complex data set, to which the statistical procedure was applied in order that this complex data set could be more comprehensible. The technique of meta-analysis could be traced back to the work of Tippet (1931), Fisher (1932), Pearson (1933), and Cochran (1937).

A number of meta-analyses had been carried out in various social and behavioral sciences, including Smith and Glass's study (1977), Mazzuca's study (1981), Willson and Putnam's study (1982), and Hillocks's studies (1984 and 1986). Experts of meta-analysis warned the learned world of the potential problems in meta-analysis when dealing with traditional literature reviews. These problems could be:

- (1) selective inclusion of studies, often based on the reviewer's own impressionistic view of the quality of the study,
 - (2) differential subjective weighting of studies in the interpretation of a set of findings,
 - (3) misleading interpretations of study findings,
 - (4) failure to examine characteristics of the studies as potential explanations for disparate or consistent results across studies, and
 - (5) failure to examine moderating variables in the relationship under examination.
- (Wolf, 1986, p. 10)

Of the three types of research defined by Glass (1976), primary research, secondary research, and meta-analysis, meta-analysis "refers to the analysis of analyses." (Glass, 1976, p. 3) Like all the new approaches that were introduced, meta-analysis was not devoid of criticisms. The major criticisms raised against meta-analysis were the following:

- (1) Logical conclusions cannot be drawn by comparing and aggregating studies that include different measuring techniques, definitions of variables (e.g., treatment, outcomes), and subjects because they are too

dissimilar.

- (2) Results of meta-analyses are uninterpretable because results from "poorly" designed studies are included along with results from "good" studies.
- (3) Published research is biased in favor of significant findings because nonsignificant findings are rarely published; this in turn leads to biased meta-analysis results.

(Glass, McGaw, and Smith, 1981).

The limitations of meta-analysis per se could include oversimplification of the results of a research domain by focusing on overall effects and downplaying mediating or interaction effects, mixing of studies that measure "apples" with those that measure "oranges" so that no meaningful results could be obtained, difficulty in treating nonindependent or multiple results within a single study, lack of studies numerous enough to insure some degree of homogeneity among these studies, and also the over-zeal in applying meta-analysis to quantitative reviews while the theoretical basis of this statistical approach was still being worked on (Wolf, 1986, pp. 53-54). There were other criticisms. However, the adherents to this approach were developing and refining meta-analysis so that the criticisms could be more or less redressed.

For all the potential weaknesses mentioned above, meta-analysis certainly has its strengths. First of all, meta-analysis is an efficient way to summarize large quantities of research literature in a specific research domain (Green

and Hall, 1984). Secondly, stronger conclusion can be reached than with impressionistic reviews. Thirdly, a sharper focus on effect can result from a large number of studies with diverse hypotheses, designs, and results. Fourth, meta-analysis does not exclude some studies as unworthy for their weak designs. Fifth, meta-analysis can be more objective than traditional literature reviews about the research quality as meta-analysis examines the effects of research quality on study findings by empirical approaches (Wolf, 1986, pp. 54-55). Furthermore, meta-analysis is very helpful in highlighting gaps in the literature, providing insight into new directions for research, and finding mediating or interactional relationships or trends either too subtle to see or that cannot be hypothesized and tested in individual studies (Green and Hall, 1984; Light and Pillemer, 1984).

In order to reduce the potential weaknesses and capitalize on the strengths of meta-analysis, Wolf listed the following guidelines for conducting a meta-analysis:

- (1) Define and report criteria for the inclusion and exclusion of studies. This is largely a substantive, conceptual issue, but quantitative issues such as how to handle a study with missing or incomplete data may also arise.
- (2) Search for unpublished studies in order to test for a type I error publication bias.
- (3) Develop coding categories to accommodate the largest proportion of the literature identified. It is better to code too many

rather than too few characteristics of studies. Both substantive and methodological characteristics need to be coded, such as type and length of the intervention, sample characteristics, research design characteristics and quality, source of study (e.g., published, dissertation, internal report, and the like), date of study, and so on.

- (4) Examine multiple independent and dependent variables separately through blocking, mediating effects. Look for interactions with the principal relationship being reviewed.
- (5) Examine and graph the distribution of results and look for outliers to examine more closely. Test for the heterogeneity of results.
- (6) Check the reliability of the raters who code the characteristics of studies.
- (7) Always accompany combined tests of significance with estimates of effect size.
- (8) Calculate both raw (unadjusted) and weighted combined tests and effect sizes to empirically examine the impact of sample size on results.
- (9) Consider whether it is important and/or practical to calculate nonparametric as well as parametric effect size estimates.
- (10) Consult the literature on meta-analysis for guidance when in doubt. Examine how others have approached particular problems in their quantitative reviews.
- (11) Combine qualitative reviewing with quantitative reviewing. Examine the comparability of treatment and control groups, respectively, from study to study. Describe interesting, worthwhile studies that are not possible to include in a quantitative review.
- (12) Describe the limitations of your review and provide guidelines for future research concerning the relationships reviewed.

- (13) Remember Green and Hall's (1984, 52) dictum:
"Data analysis is an aid to thought, not a
substitute." (Wolf, 1986, pp. 55-56)

Literature Reviews and Critiques

By nature, meta-analysis was the application of statistical procedures to the summary of large numbers of research reports. The literature reviews were the commonest practice in research projects to gather information about what had been done in the area of research interest, what trends and issues existed there, and what possible traps there were to be cautious about. This type of literature reviews served a number of purposes. Traditionally, literature reviews were done in a non-statistical approach, dwelling heavily on the individual researcher for significance, direction, and inspiration. In the research area of writing and teaching writing, Braddock, Lloyd-Jones, and Schoer's report (1963) was an example of traditional literature review; and Hillocks's report (1984) was an example of meta-analysis. Usually, the traditional literature reviews dealt with a much smaller number of research reports than meta-analysis. In the selected research reports for analysis in this study, Braddock, Lloyd-Jones, and Schoer's report(RA); Cooper's report (RDC); Stotsky's report (RI); and Söter's report (RMS) were the four examples of traditional literature reviews (Table 21).

Table 21

TRADITIONAL LITERATURE REVIEWS

Code	Year	Author(s)	Theme
RA	1963	Braddock, Lloyd-Jones & Schoer	Experimental studies
RDC	1975	Cooper	Measurements of growth
RI	1983	Stotsky	Read/Write relationships
RMS	1987	Söter	Writing across curriculum

The scope and criteria of selection of research reports to be included in these studies were of different types. RA used a committee of experts to decide which reports should be included in their study. The purpose of their study was to define the direction and properties desired for future research studies in the field of teaching writing. To achieve this purpose, a content analysis in the form of detailed summary was conducted on the five empirical studies selected.

This process was characterized by the formation of an ad hoc committee appointed by the Executive Committee of the National Council of Teachers of English based on concern over the public pronouncements about how writing should be taught, obtained grants, clarified purposes, wide search of existing literature, expert consultation, discussion of concerns about the report, writing of the report, acquisition of feedback from original authors, distribution of reports to members of the ad hoc committee and the NCTE

Committee on Publications and the final publication of the report (Braddock, Lloyd-Jones, and Schoer, 1963, pp. 1-4). Based on the extensive involvement in this project, the researchers not only finished the selection of the reports for study, made a detailed summary of these reports, but also identified both the research areas where some research had been carried out and the research areas where research was needed.

Areas where some research had been carried out included the following:

The Writer Variable:

Kincaid (1953)
Miller (1958)
Anderson (1960)

The Assignment Variable:

Kincaid (1953)
Wiseman and Wrigley (1958)

The Rater Variable: (no citation)

The Colleague Variable:

Smith (1931)
Stalnaker (1934)
Hartog & Rhodes (1936)
Cast (1940)
Wiseman (1949)
Starring (1952)
Harris (1962)

Frequency Counts:

Thorndike (1927)
Leonard (1932)
LaBrant (1933)
Anderson (1937)
Evans (1939)
Boder (1940)
Chotlos (1944)
Van Bruggen (1946)
Kaldegg (1950)
Kimoto (1951)

Kincaid (1953)
Wiseman and Wrigley (1958)
Haskins (1960)
Matheson (1960)
Strickland (1962)

Research Methodology:

Rivlin (1938)
Dolch (1956)
Carroll (1961)
Singleton (1961)

Areas where research was needed were identified as follows:

1. What kinds of situations and assignments at various levels of schooling stimulate a desire to write well?
2. What do different kinds of students prefer to write about when relieved of the expectations and requirements of teachers and others?
3. What are the sources of fear and resentment of writing?
4. How do the kinds of writing which adults compose vary with their occupations and other factors?
5. What is the effect on writing of having the student compose his paper for different kinds of readers?
6. At which levels of maturation does it seem appropriate to introduce the various modes of discourse--narration, poetry, drama, exposition, argument, and criticism?
7. What is the relative effectiveness of writing shorter and longer papers at various levels of maturity and proficiency?
8. At which levels of maturation does it seem appropriate to introduce the various rhetorical elements of writing?
9. What are the effects of various kinds and amounts of reading on the quality and kinds of writing a person does?

10. What are the direct and indirect effects of particular sensory experiences and guided observation upon writing?
 11. At what stages of maturity do students spontaneously seek specific help in improving particular aspects of writing, such as specificity of details, transitions, parallel structure, and metaphor?
 12. At which levels of maturation can particular aspects of writing most efficiently be learned?
 13. Does the oral reading of rough drafts help the elementary school child strengthen "sentence sense"? How does it?
 14. What techniques of composition most effectively help build self-discipline and pride in clarity, originality, and good form?
 15. What procedures of teaching and learning composition are most effective for pupils of low socioeconomic patterns?
 16. What procedures of teaching and learning composition are most effective for pupils learning to write English as a second language?
 17. Can study of the newer types of linguistics help writers?
 18. Can formal study of rhetorical theory or of logic help writers?
 19. How is writing affected by extensive study and imitation or parody of models?
 20. What forms of discourse have the greatest effect on other types of writing? For example, does writing poetry help a writer of reports?
 21. What is involved in the act of writing?
 22. How does a person go about starting a paper? What questions must he answer for himself?
 23. How does a writer generate sentences?
 24. Of what does skill in writing really consist?
- (Braddock, Lloyd-Jones, and Schoer, 1963, pp. 52-53)

Cooper's research review (RD1, 1975) focused on the issue of measuring growth in writing, with a view to establishing a feasible writing program for school. Cooper began by differentiating between the definition of measuring and the definition of evaluating, the definition of formative measurement and the definition of summative measurement, the definition of criterion-referenced measurement and the definition of norm-referenced measurement. Cooper put emphasis on the behavioral objectives of the writing curriculum. In this light, Cooper also commented on Foley's table of specifications for writing instruction (1971).

The core part of RD1 was a set of objectives and measures applied to the implementation of the objectives, which was summarized as follows:

Objective (O) and Measurement (M)

- O1. Each student will make fewer deviances from standard written usage.
- M1. Error counts in 300 written words from two pieces of writing both at the beginning and at the end of the instructional period.
- O2. Each student will achieve greater syntactic fluency.
- M2. T-unit length of forty-five T-units from at least three pieces both at the beginning and at the end of the instructional period.
- O3. Each student will demonstrate growth in the ability to write an acceptable piece in each of the following kinds of writing:
 - journal entry
 - personal letter
 - autobiographical incident
 - autobiographical phase

- biography
- chronicle
- entry in a log of response to literature
- M3. Multiple and complex measure by self, peer, and teacher of several pieces of writings both at the beginning and at the end of the instructional period.
- O4. Each student will show a willingness to write.
- M4. Teacher's observation and the accumulation of papers in student's folder.
- O5. Each student will come to value writing and writers.
- M5. A Likert-type attitude measure, teacher's recorded observations, and a statement from the student.
- O6. Each student will be willing to make a contribution to his small-group writing workshop.
- M6. Teacher's observations and self-evaluation by student and group.

Based on the above objectives and relevant measurement, Cooper offered the following cooperative whole-school measurement scheme based on research finding:

Getting Ready:

1. Faculty contribute topics, limit the number to twenty by consensus. Chair picks four for the four writing days of the school.
2. Chair makes ID slips for student writers, four for each; and copies of Reader's Analytic Scale, two for each ID slip.
3. Teachers practice using the Analytic Scale.
4. Teachers agree on four school writing days, two at the beginning, and two at the end of the semester.

On Writing Day:

1. Chair announces the topic.
2. Students write on the topic.
3. Teachers collect the papers with numerical ID.
4. Reorder the papers and randomly distributed to teachers.

The Rating Process:

1. Teachers rate the papers on the Scale, using a reader number.
2. Teachers return the first set of papers and take a second set.
3. Chair's office puts the papers, slips, and scales together by ID.
4. Office enters the composite score of all four papers in record file.

Using the Results:

The score could be used in a number of ways, such as identifying individual student's ranking in grade or in school for placement purposes, comparing strengths and weaknesses of writing programs, examining writing abilities of different groups of students, investigating strengths and weaknesses in instructional implementation, and so on.

(Cooper, 1975, pp. 118-119)

Stotsky's study (1983a) summarized the research studies on reading/writing relationships and suggested new directions for future studies. According to Stotsky, studies investigating the writing/reading relationships fell into three categories: 1. correlational studies, 2. studies of influence of reading upon writing, and 3. studies of influence of writing upon reading (Table 22).

After the extensive survey, Stotsky offered the following list of suggestions for future research:

1. Research is needed on the traits of good readers/poor writers or poor readers/good writers.
2. Research is needed to develop instruments that systematically relate reading to writing.

Table 22

STUDIES OF WRITING/READING RELATIONSHIPS

Category	Researcher Year	Major Findings
Correlational Studies	Schonell 1942	A high correlation exists between measures of composition and measures of vocabulary, etc.
	Piexotto 1946	A low, significant correlation exists between scores on English Essay Test, reading test, & SAT verbal section.
	Diederich 1957	A higher correlation exists between teachers' composition score and SAT verbal score.
	Monk 1958	Good writers read more in leisure time.
	Loban 1963	A high correlation exists between reading and writing quality.
	Fishco 1966	A significant correlation exists between creative writing and reading comprehension.
	Donelson 1967	Effective writers tend to be readers.
	Maloney 1967	Superior writers read more and better and tend to be female.
	Woodward 1967	Toor freshman writers read less in high school.
	Barbig 1968	Good writers do more voluntary reading and tend to be female.

Table 22 (continued)

STUDIES OF WRITING/READING RELATIONSHIPS

Category	Researcher Year	Major Findings
	LaCampagne 1968	Superior writers read more.
	Woodfin 1968	Best predictor for writing is reading ability and language score.
	Zeman 1969	Reading comprehension grows with use of compound & complex sentences.
	Grimmer 1970	A significant correlation exists between composition quality and reading achievement.
	Calhoun 1971	A significant correlation exists between writing achievement and reading comprehension.
	Siedow 1973	No correlations exist between cloze test score and syntactic maturity of rewritten kernel sentences.
	Evanechko 1974	Significant correlations exist between reading achievement, fluency, and syntactic complexity.
	Fuller 1974	No difference exists in T-unit between good and poor readers.
	Heil 1976	A significant correlation exists between T-unit and comprehension.
	Lazdowski 1976	Proficiency in writing reflects proficiency in reading.

Table 22 (continued)

STUDIES OF WRITING/READING RELATIONSHIPS

Category	Researcher Year	Major Findings
	Thomas 1976	A significant correlation exists between writing achievement and reading comprehension.
	Bippus 1977	A significant correlation exists between writing productivity and reading comprehension.
	D'Angelo 1977	A significant correlation exists between reading scores and writing achievement.
	Grobe 1977	Good writers have higher reading scores.
	Perron 1977	Low, medium, high reading ability groups write at significantly different levels of syntactic complexity.
	Evans 1979	An inverse correlation exists between cloze test score and syntactic complexity by rewriting kernel sentences.
	Heller 1979 /1980	Better readers have longer T-unit length.
	Felland 1980	Superior writers read more from a survey.
	Johnson 1980 /1981	Significant correlations exist between writing and vocabulary more than comprehension.

Table 22 (continued)

STUDIES OF WRITING/READING RELATIONSHIPS

Category	Researcher	Year	Major Findings
Influence of Writing upon Reading	Shanahan	1980	Reading and writing are related in different ways in different grades.
	Atwell	1981	Better writers reread and plan more in writing.
	Birnbaum	1981	Patterns exist of high quality writing and reading behavior with related proficiency.
	Taylor	1981	Reading comprehension, not listening comprehension, is related to course grade.
	Barton	1930	Outlining improves retention of information in content areas.
	Newlun	1930	Summary writing improves mastery of information in history.
	Dynes	1932	Writing is superior to rereading in retention of information.
	Salisbury	1934	Outlining and summary writing improves reading test.
	Jencke	1935	Some superiority exists for précis on vocabulary and untimed reading interpretation test.
	Smith	1971	No gain exists for "creative" or "noncreative" writing on short comprehension test

Table 22 (continued)

STUDIES OF WRITING/READING RELATIONSHIPS

Category	Researcher Year	Major Findings
	Obenchain 1971	Writing program improved writing at .06 level.
	Oehlkers 1971	One-year creative writing shows no improvement in word recognition in 1st graders.
	Nagle 1972	Directed writing program improved general reading comprehension, but not in content areas.
	Kulhavy 1975	Note-taking is better than underlining in performance.
	Wolfe 1975	Writing new words in sentence is better than writing new words alone.
	Pederson 1977	Sentence-combining affects "lexical density"; results are difficult to interpret.
	Doctorow 1978	One sentence summary writing improves comprehension and recall.
	Taylor 1978	Paraphrasing and summarizing improves summarizing.
	Collins 1979	Expressive writing with reading improves reading comprehension.
	Combs 1979	Sentence-combining has ambiguous effect on reading comprehension in a synthesis study.
	Bretzing 1979	Note-taking with paraphrasing and summarizing is better than note-taking verbatim in recalling.

Table 22 (continued)

STUDIES OF WRITING/READING RELATIONSHIPS

Category	Researcher	Year	Major Findings
Influence of Reading upon Writing	Taylor	1980	One sentence summary writing improves comprehension and recall in social study.
	Glover	1981	Paraphrasing is better than key word writing in recall.
	Walker-Lewis	1981	Expressive and other writing with reading improves reading comprehension.
	Eurich	1931	No gains exist in reading or writing after a reading program.
	Clark	1935	Additional reading improves grammar and usage.
	Bagley	1937	Literature readers write better than grammar readers.
	Mathews	1945	Poor writers did not gain by reading program.
	Heys	1962	Students reading more write better than students writing more.
	Christiansen	1965	Reading more compensates for writing more in writing improvement.
	DeVries	1970	Students who read instead of writing write better.
	Schneider	1971	Remedial writers did not gain by reading treatment.
	Miller	1974	Remedial writers did not gain by reading program.

Table 22 (continued)

STUDIES OF WRITING/READING RELATIONSHIPS

Category	Researcher	Year	Major Findings
	Mills	1974	Children involved in children's literature write better.
	Andreach	1975	Literature study improves composition organization.
	Campbell	1976	No significant difference exists after a reading program treatment.
	Bossone	1976	Better expository writing follows a reading program.
	Elley	1976	Readers of literature write as well as readers of grammars.
	Maat	1977	No gains exist in writing after a reading program.
	Belanger	1978	No gains exist in writing after a reading program.

Source: adapted from Stotsky, 1983a.

3. Research is needed to describe the developmental phase of reading/writing relationship.
4. Research is needed in case study for a clearer picture of the reading/writing relationship.
5. Research is needed to develop a better measure for lexical growth.
6. Research is needed to find out the reading behavior of good writers and poor writers in reading others' as well as their own writing.
7. Research is needed in ESL on the above issues.
8. Research is needed to develop a theoretical framework for this field of research.

(Stotsky, 1983, pp. 636-637)

Söter's study (1987) summarized the research studies on writing across the curriculum and suggested directions for future studies. She discussed a number of issues of public interest, dwelling on the definition of school writing, writing process, assignment variable, cognitive involvement, textual and contextual perspectives, relationship between reading and writing, sociocultural perspective, developmental perspective, instructional perspective, and writing evaluation (Table 23).

Table 23

DIMENSIONS IN WRITING ACROSS THE CURRICULUM RESEARCH

Dimension	Research	Year	Concerns
Definition of School Writing	Reed Reed Graves Bissex	1971 1975 1975 1980	writing involves many levels of cognitive performance: in form, function, and content; writing is a social function; content-area writing is hindered by early demand on written convention.
Process of Writing	Emig Flower Flower Perl Britton	1971 1977 1979 1979 1982	process involves prewriting activities like brainstorming, note-taking, list-making, planning; writing activities like drafting, revising, writing; post-drafting activities or editing
Assignment Variable	Crowhurst Applebee Baker Hillocks Hoetker Purves	1979 1981 1982 1982 1982 1986	several constraints exists: knowledge: drawing information from memory, integrating old\new information, discarding information, determining on relevance. written language: usage, construction of sentence. self-monitoring: ability to direct process, ability to arrange information. rhetorical\sociocultural\ethical: purpose, audience, role in product.
Textual & Contextual	Britton Olson Tannen Cooper Purves	1975 1977 1982 1984 1986	compared with oral language, written language must have: identity of writer, established context, clarified object for discussion, choice of lexis and punctuation, an imagined audience, well-defined structure.

Table 23 (continued)

DIMENSIONS IN WRITING ACROSS THE CURRICULUM RESEARCH

Dimension	Research	Year	Concerns
Reading & Writing	Stotsky	1984	Reading and writing are: correlated, highly continuous, recurring, and recursive, involved with cognitive process.
Sociocultural Concerns	Cole Ong Langer Purves	1974 1982 1984 1986	Cultural context operates in: how we think, how we use discourse structure, how we use linguistic resources, how we write on a topic, what knowledge we have, what assignment we give, what role we take with audience, what purpose we have in writing.
Developmental Concerns	Graves King Dyson Bereiter Purves	1975 1979 1982 1983 1986	Four major aspects of writing: graphic/mechanic, linguistic, semantic/textual, rhetorical. Vital factors in language development: sense of meaning, sense of convention, sense of audience, sense of development.
Instructional & Evaluative Concerns	Olson Applebee Beach Carroll	1981 1981 1984 1984	Problems: psychologist vs. educationist, anthropologist vs. educationist, consensus vs. reality, theory vs. practice, standardized testing vs. pupil, information vs. writing skill.

Source: Söter, 1987.

Research Critiques and Reflective Study

In contrast to the comprehensive literature reviews, some literature discussions were focused on one or two issues. The researcher could carry out more in-depth discussions based on personal opinions, other sets of data leading to different conclusions, or different perspective on the same set of data. Therefore, different conclusions could be drawn. This type of research reports was termed research critique or reflective reports. This study included two reflective studies (Table 24).

Table 24

REFLECTIVE STUDIES

CODE	YEAR	AUTHOR	THEME
RML	1987	Lutkus	T-unit
RN	1988	Reed	Instruments

The Lutkus's study (1987) compared the T-unit measure for syntactic maturity developed by Hunt (1965) and the sentence weight measure developed by DiStefano and Howie (1979) by using Hunt's "Aluminum" test package. Hunt found young writers moving through three stages of coordination. Fourth graders normally coordinated main clauses. Children between 4th grade and sixth grade learned to coordinate predicates. Beginning from sixth grade, children learned to coordinate objects. This process reflected itself in the

uneven increase of the T-unit measurement, but not in the sentence weight measurement. Sentence weight measurement was a measure of modification, calculated by attributing a score of 1, 2, 3, or larger numbers to the basic words, their modifiers, the modifiers of the modifiers, and so on. Then the scores were added together to be divided by the total number of words, not counting the structure word, like preposition, articles, and the like. With the simple sentences where fourth graders to sixth graders were developing their coordinating abilities, the words were without modifiers, resulting in a constant 1 as the sentence weight measure. Therefore, Lutkus argued that sentence weight measure was less sensitive than T-unit measure.

When the coordination was neutralized by adding the missing words, T-unit measure was found to be a better indicator than sentence weight measure. Although coordination was not a major sign of syntactic maturation, for 4th to 6th graders, coordination was a major change (Hunt, in Lutkus, 1987, p. 50). A more relevant criticism was that the sentence weight measure was more difficult to calculate, the reason being that raters of different grammatical orientations might argue about the score to be assigned to a certain word, whereas for T-unit measure, it was not necessary to differentiate the functions of each word.

In another research critique, Reed, Burton, and Vandett

(1988) discussed two of the most frequently used measuring instruments by writing researchers, one being Daly and Miller's Writing Apprehension Test (1975) and the other being Hunt's T-unit Analysis (1970). The chief concern about Daly and Miller's Test was related to the adoption of the 3-uncertain response choice. The researchers challenged the rationale of the inclusion of this choice in the test. In a study of 1984, the researchers applied the Daly and Miller Test to a group of 83 high school seniors with the Missouri College English Test (Form B) to determine the relationship between writing skills and writing apprehension. In this study, 42 of the subjects chose "uncertain" to seven or more of the 26 statements (Reed, Burton, and Vandett, 1988, p. 2). The researchers made a data modification by forcing a directional choice based on the total score. Therefore, all the uncertain scores were changed to 2 if the total score was below average, and all the uncertain scores were changed to 4 if the total score was above average. The reason why this approach was adopted was based on the Nunnally study (1967), where it was pointed out that a response style of cautiousness was possible when the correct answer was in doubt. It was a reflective of personality, rather than an inability to answer. The adoption of 3-uncertain practically gave the subjects the choice of not giving their true scores. As a consequence of this technique, the results made the difference between

significant and insignificant findings (Table 25).

Table 25

DIFFERENCE BETWEEN ORIGINAL DATA AND MODIFIED DATA
IN REED, BURTON, AND VANDETT'S STUDY

Variable	Beta	SEB	B-weight	F	p
Original Data*					
Skills	-.58	.21	-.31	7.84	.007
Gender	1.63	3.89	.05	.17	.683
Prep.	1.56	2.77	.06	.32	.575
Constant	90.93	10.18	--	8.93	.000
Modified Data**					
Skills	-.73	.27	-.31	7.20	.009
Gender	2.10	5.20	.05	.16	.687
Prep.	.97	3.62	.03	.63	.789
Constant	96.53	13.13	--	7.25	.000

* [Overall Regression] $F(3,79)=2.80, p<.05$

** [Overall Regression] $f(3.79)=2.49, p>.05$

Source: adapted from Reed, Burton, and Vandett, p.2.

A second concern was about the relationship between quality of writing and the T-unit syntactic complexity measure. The researchers challenged the rationale of using T-unit measure as an index for quality of writing. The researchers contended that no research up to then was conclusive on the relationship between writing quality and T-unit measure. Two studies conducted by Reed and Sherman (1991) and Reed and Vandett (1986) indicated either lack of relationship or a negative relationship between quality and

syntactic complexity.

Reed and Sherman's study (1991) used a 2 X 2 factorial design on the variables of external storage and heuristic prewriting activities. Four classes of basic writers and four classes of honors writers were randomly assigned to the four groups. However, the results showed that all the honors writers groups and two basic writing groups scored significantly higher on words per T-unit (Table 26).

Table 26

MEAN SCORES ON STUDENT ESSAY FOR HOLISTIC QUALITY
AND WORDS PER T-UNIT,
BY EXPERIMENTAL CONDITIONS OF ABILITY,
EXTERNAL STORAGE, AND HEURISTIC

Writing Ability	Descriptor	Treatment			
		ES+H+	ES+H-	ES-H+	ES-H-
Basic	Quality	3.50+	3.33+	2.50+	3.91+
	W/T-unit	13.67	12.83*	11.78*	17.61*
Honors	Quality	6.43*+	5.96+	6.09+	5.43*+
	W/T-unit	15.38	14.63	13.71	13.89

* p<.05 within writing ability groups

+ p<.05 between writing ability groups

Source: adapted from Reed, Burton, and Vandett, 1988, p. 5.

In another study by Reed and Vandett (1986), the quality and the syntactic complexity of essays by basic writers were analyzed. The essays either had a group experience as topic or a personal experience as topic.

Again, the personal experience group scored higher on quality, yet the group experience group scored higher on words per clause measurement (Table 27).

Table 27

MEANS AND STANDARD DEVIATION (IN PARENTHESES)
OF HOLISTIC QUALITY SCORES AND
WORDS PER CLAUSE FOR TOPIC-TYPE (GROUP OR INDIVIDUAL)

Dependent Measure	Topic Type	
	Group	Individual
	n=10	n=33
Quality*	2.70 (1.06)	3.67 (1.02)
Words/Clause**	9.74 (2.44)	8.60 (1.44)

* $F(1,41)=6.771$, $MSe=1.059$, $p=.012$

** $F(1,41)=3.427$, $MSe=2.926$, $p=.06$

Source: Reed, Burton, and Vandett, p. 6.

Based on the above results, the researchers contended that T-unit measure as a measure for syntactic maturity could not be equated to be a measure for quality (7).

Hillocks, Jr.'s Study and the Meta-Analysis

Hillock, Jr.'s study (1984) utilized the statistic technique known as the meta-analysis and summarized the research findings of about five hundred empirical studies carried out after Braddock, Lloyd-Jones, and Schoer's study of 1963. This screening process was governed by the

following minimum criteria:

1. The study had to involve a treatment;
2. The study had to use a scale of writing quality applied to samples of writing;
3. The study had to exercise minimal control of teacher bias;
4. The study had to control for differences among groups of students;
5. Compositions had to be scored under conditions that help to assure validity and reliability.

(Hillocks, Jr., 1984, pp. 135-136)

The techniques adopted by the researcher were developed by Glass (1978). The model adopted for this study was developed by Hedges (1981, 1982a). As defined by Hillocks, Jr., a meta-analysis is a statistical procedure which "computes standard scores for various treatments' gains or losses by dividing the difference between post-test scores, adjusted for the difference between pre-test scores, by the pooled standard deviation of post-test scores for all groups in the study." (p. 136) Thus, if a given treatment was said to have an experimental/control effect size of .5 standard deviations, it meant that the gain for the average student in the experimental group is .05 standard deviations greater than for the average student in the control group (p. 136). This study also used the technique developed by Hedge which weights each effect size by the reciprocal of its variance so that the accumulation is not simply an average of raw effect sizes, but a mean effect size dependent on the

variance of its constituents (p. 136).

In order to explain the variability among the characteristics of the treatments in relation to the variability of their effect sizes, the researcher took the following steps: 1. categorizing the treatments along various dimensions: such as instructional mode, focus of instruction, and duration; 2. comparing the mean effect sizes of treatments grouped together; and 3. testing the studies grouped together for homogeneity.

Two independent coders with an intercoder reliability of .87 coded all the variables in the selected studies. The differences were resolved by conference. The researcher agreed with .98 of the coders' resolved coding with the differences settled by consultation. With the sophisticated coding system, it was possible to examine several dimensions. The report concentrated on three dimensions: duration of treatment, mode of instruction, and focus of prewriting instruction (p. 138). By cross tabulation, treatments could be grouped by different dimensions.

After the overall mean effect sizes and homogeneity statistics were established, four studies that contributed most to heterogeneity were removed, two with high positive effect sizes and two with high negative effect sizes. At most, 15% of the treatment groups could be removed to achieve homogeneity. If homogeneity could not be achieved, it must be concluded that the dimensions and the categories

did not have a high level of an explanatory power (p. 138).

After the techniques were applied to the different dimensions, the results were as follows (Table 28):

In conclusion, this chapter discussed the different types of research literature summaries, reviews, research critiques, and meta-analysis used in the published reports selected for this study. Some of the strengths and weaknesses of the various types of secondary research were also discussed. Besides, some interesting research findings and suggestions for future research were listed in this chapter for further reference.

Table 28

RESULTS OF META-ANALYSIS BY DIFFERENT DIMENSIONS

Dimension	n	Mean Effect	SD	H	Homogeneity p<.01
All	75	.28	.018	411.08	-
Mode of Instruction	29	.24	.019	73.83	-
Natural Process	9	.19	.037	23.15	-
Environmental Mode	10	.44	.050	12.83	+
Presentational Mode	4	.02	.114	.92	+
Individualized Mode	6	.17	.064	14.68	+
Categorized by Mode of Instruction	29			51.58	-
Focus of Instruction	39	.26	.023	84.48	-
Grammar	5	-.29	.059	8.85	+
Sentence Combining	5	.35	.083	1.89	+
Models	7	.22	.057	5.31	+
Scales	6	.36	.078	6.89	+
Free Writing	10	.16	.035	27.25	-
Inquiry	6	.56	.076	8.73	-
Categorized by Focus of Instruction	39			58.92	-
(Modified)				49.90	+
Duration					
under 13 weeks	39	.28	.028	153.48	-
over 12 weeks	34	.28	.024	277.61	-
under 17 weeks	58	.27	.021	395.74	-
over 16 weeks	15	.31	.038	34.48	-

Source: adapted from Hillocks, Jr., 1984.

CHAPTER VI

DISCUSSION AND CONCLUSION

As the concluding chapter, this chapter will address two important issues involved in the whole study: (1) the hierarchy theory as can be used in this scenario, and (2) the program outline dedicated to this researcher's hard-working Chinese colleagues. To achieve this purpose, this researcher will refer to the previous chapters every now and then.

From the previous chapters, it has become clear that, according to research in the developmental studies, writing as learning is to be acquired through a well designed curriculum, where the trained instructor teaches to mastery in the light of the goals set in the curriculum. Writing as human behavior can be broken down to various discreet technical components, which can be effectively taught to the students by means of well designed classes. The success of a composition course has to be defined in the light of the goals set for the composition class. Unintentionally acquired writing techniques, like windfalls, are not necessarily the results of classroom instruction.

Cognitive research findings indicate that the writing techniques to be taught to the students have to be at the right level of difficulty for each grade or class where writing is actually taught. Each writing technique can be addressed at different levels of difficulty from the

cognitive perspective.

Studies in the variable of assignment indicates that the existing writing assignments strongly suggest a liberal art approach to writing instruction. The functional approach is therefore suggested for composition classes where there is a manifested need for the writing in various functions in the school district and community. It will show the students the real need for writing.

The fact that these studies were carried out many years ago, and that they had made their contributions to the development of this branch of learning is self evident. It is most important today to improve this field of learning based on the development of modern technology and progress in other related fields of study as well as in the field of writing and teaching writing by adopting a more meaningful theoretical structure.

If we examine the progress in the physical sciences, we can see that the branches of studies in those fields are gradually merging into one another, and a formidable organized mass is looming up. Human behavior is an even more formidable mass with an organized and structured nature. The difficulty confronting us is to envisage the instrument to observe and measure the kaleidoscope of human behavior, including the behavior of writing and teaching writing. The human writing behavior is an extremely complex process. All the variables in the imaginable hierarchy and

all the imaginable related sub-hierarchies exercise some influence on the process of writing in all kinds of imaginable manners, so that the existing statistical processes designed for measuring and testing individual factors seem inadequate to come up with a satisfactory result. No correlation of 1.0 can be detected in the human behavior, simply because no single variable can be isolated and no perfect control condition can be achieved.

The hierarchy theory will provide such a structure. When we reflect on the concepts to be tested in the various studies, we can detect the hierarchy structure embedded there. For example, the concept of "method" is actually a sub-hierarchy including subject-environment-material-instruction-teacher-cognition-learning style-motivation-commitment-and-what-you-name-it. When we look at each layer in this sub-hierarchy, each layer displays an independent small subhierarchy. For example, the subhierarchy of subject includes physical development-physiological difference-mental development-individual learning style-language development-attitude toward learning-maturity level-value system-cultural background-and-what-not.

Basically, the human writing behavior involves the following subhierarchies in a meaningful pattern:

1. dimension of cognition:

perceptive process

cognitive process

- developmental process
- affective process
- 2. dimension of communication
 - coding process
 - decoding process
 - interpreting process
 - transmitting process
- 3. dimension of environment
 - functional process
 - adaptive process
 - nurturing process
- 4. dimension of culture
 - assimilating process
 - differentiating process
 - valuative process
- 5. dimension of schooling
 - maturing process
 - learning process
 - interactive process
- 6. dimension of information
 - input process
 - storing process
 - retrieving process
 - output process
- 6. dimension of language
 - concept process

verbalizing process
transformational process
conformity process
rhetoric process

There might be many other dimensions in which the writing activity is involved. This list shows that the understanding of the human writing behavior is far from organized and structured. Firstly, these processes are actually subhierarchies. Each of them contains a number of subprocesses. This process can go on to infinity. Secondly, these dimensions plus the dimension of time and the dimension of space are interrelated and will interact upon one another. Thirdly, since each dimension, or process, of even a subprocess will require a lifetime to study only to reach a certain depth, it is not practical to ask anybody to know all those dimensions. Fourthly, a certain amount of ambiguity has to be allowed in dealing with issues involved in the human writing behavior. Therefore, for the above frustrations, the theory of hierarchy can provide some insightful solutions:

1. The hierarchy theory offers a theoretical framework for further research direction. Research questions can be defined in a more organized way.

2. The hierarchy theory offers a direction into which enquiring professionals can search for related information and revelations.

3. The hierarchy theory can identify the level of sophistication at which problems occur and at which solutions seem to be accessible.

4. The hierarchy theory offers each specific area of study a macrocosmic framework and at the same time delineate the relationship of each field of study to other fields.

5. The hierarchy theory combines research fields in a meaningful unity and relates it to the basic functions of the human brain, which operates in an identified hierarchical pattern.

6. The hierarchy theory has the potential of being developed in either vertical or horizontal directions.

7. The hierarchy theory will promote a global view in dealing with controversial issues.

8. The hierarchy theory will unite all available resources for research in a more organized way, a more effective way, and a more efficient way.

Since this dissertation is meant to be a dedication to the author's colleagues in China, who are intelligent, hard-working, committed, yet severely limited in their access to available information, it is desirable here to offer a program for the initiation and development of research in teaching English composition in China.

To promote research in the teaching of English composition in China, it is necessary that a national organization be established, or this function be attributed

to one of the existing academic associations through a newly established panel. A number of universities can be engaged to start the procedures, such as the collection of documents, the drafting of guidelines, the identification of research questions, the training of researchers, the location of fund, the designing of goals, the planning for surveys, the establishment of norms, the preparation of facilities and equipment, the compilation of textbooks, and the publication of a journal for this specific area. A center can be established to conduct this enterprise and coordinate the activities, such as providing a one-year training program in research methodology and statistics study for English major graduates from all over the country, and inviting experts from the United States to conduct summer seminars on various topics identified as crucial to the development of such programs.

As literature review for social and behavioral research tends to be subjective and impressionistic (Wolf, 11), it would be desirable to incorporate research courses as part of a standard curriculum into graduate programs, at least tentatively, introducing various types of experimental designs (Campbell and Stanley, 1963), access to the body of research studies (Glass, McGaw, and Smith, 1981; Hunter, Schmidt, and Jackson, 1982) through computer search in an identified field, and basics of meta-analytical techniques (Glass, McGaw, and Smith, 1981) to integrate the effects of

previous studies so that this approach could be built into any research format, including dissertations.

However, there are some particular concerns in the implementation of the guidelines discussed here into a Chinese context. First of all, educational policy is made by the central government and the central committee of the Communist Party of China. Local schools have all the party and government stipulations, policies, and documents to observe, which have been known to be rigorous and frequently contradictory, reflecting the constant conflict in the decision-making process at the national level. Most probably, the purpose of the research will be restricted to the issues of classroom methodologies. Anything that is somehow related to the management and governance of the school, the personnel, and the funding, or issues related to governmental policies and stipulations will most probably be ignored at the best, and persecuted at the worst, unless these efforts are attempting to consolidate the "party leadership" in school. The argument of "who controls the schools" has been notorious among Chinese professionals. Anybody who shows any sign of deviation from the party's political line has been identified as unfit for administrative positions in school. Job stability makes it even harder for faculty members to criticize the current practice of the administration, even in a constructive way, as frequently the faculty members are the ones who suffer in

their conflict with the administrators, as the administrators are in control of almost everything that is directly related to the daily work and life of the faculty members. It is absolutely necessary to improve this situation before any conscientious research effort can be made fruitful.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Adipattaranun, N. (1992). An examination of the variables in the writing process of ESL/EFL students in a process-oriented freshman composition course (Doctoral dissertation, Indiana University, 1992). Dissertation Abstracts International, 53, 1826A.
- Agnew, N. M., & Pyke, S. W. (1978). The science game: An introduction to research in the behavioral science. Englewood Cliffs, NJ: Prentice-Hall.
- Allameh, E. J. G. (1990). An investigation of trends and emerging issues in foreign student enrollment in United States higher education with major emphasis on language concerns and a case study comparing the performance of native students with nonnative students on the test of English as a foreign language (Doctoral dissertation, University of Kentucky, 1990). Dissertation Abstracts International, 51, 1126A. Allameh, E.
- Anastasi, Anne. (1968). Psychological testing (3rd ed.). New York: Macmillan.
- Anderson, C. C. (1960). The new STEP Essay Test as a measure of composition ability. Educational and Psychological Measurement, 20, 95-102.
- Anderson, J. E. (1937). An evaluation of various indices of linguistic development. Child Development, 8, 62-68.
- Anderson, R. C. (1977). The notion of schemata and the educational enterprise. In P. C. Anderson, P. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge (pp. 415-431). Hillsdale, NJ: Erlbaum.
- Andreach, J. (1975). The use of models to improve organizational techniques in writing (Doctoral dissertation, Yeshiva University, 1975). Dissertation Abstracts International, 36, 4980A.
- Applebee, A. N. (1981). Writing in the secondary school: English and content areas (Report No. 21). Urbana, IL: National Council of Teachers of English.
- Applebee, A. N. (1982). Writing and learning in school setting. In M. Nystrand (Ed.), What writers know (pp. 365-381). New York: Academic Press.

- Applebee, A. N. (1984). Contexts for learning to write: Studies of secondary school instruction. Norwood, NJ: Ablex.
- Applebee, A. N., Langer, J. A., & Mullis, I. V. (1986). The writing report card: Writing achievement in American schools. Princeton, NJ: Educational Testing Services.
- Aristotle. (1899). Politics (B. Jowett, Trans.). New York: Colonial Press.
- Arnold, C. L. (1992). Methods, plainly speaking: An introduction to hierarchical linear models. Measurement and Evaluation in Counseling and Evaluation, 25, 58-90.
- Atwell, M. (1981, November). The evolution of text: The interrelationship of reading and writing in the composing process. Paper presented at the Annual Conference of the National Council of Teachers of English, Boston, MA.
- Babbie, E. R. (1973). Survey research methods. Belmont, CA: Wadsworth Publishing Company, Inc.
- Bacon, F. (1889). Advancement of learning and novum organum. New York: Colonial Press.
- Bagley, D. (1937). A critical study of objective estimates in the teaching of English. British Journal of Educational Psychology, 7, 57-71; 138-155.
- Bailey, D. (Ed.). (1965). Essays on rhetoric. New York: Oxford University Press.
- Bain, A. (1867). English composition and rhetoric (2nd English ed.). New York: D. Appleton and Company.
- Baker, E. (1982). The specification of writing tasks. In A. C. Purves, & S. Takala (Eds.), An international perspective on the evaluation of written composition. Evaluation in Education: An International Review Series, 5, 205-390.
- Bamberg, B. (1983). What makes a text coherent? College Composition and Communication, 34, 417-429.
- Banister, L. (1991, March). The feminist rhetorics of Janet Emig and Andrea Lunsford. Paper presented at the meeting of the Conference of College Composition and Communication, Boston, MA.

- Barbig, E. V. (1968). An exploration of growth in written composition to determine the relationship of selected variables to poor writing in grades 9 and 12 (Doctoral dissertation, University of Tennessee, 1968). Dissertation Abstracts International, 29, 2140A.
- Bartlett, F. C. (1932). Remembering: A study in experimental and social psychology. Cambridge: Cambridge University Press.
- Barton, W. (1930). Outlining as a study procedure. Contribution to education (Report No. 411). New York: Bureau of Publications, Teachers College, Columbia University.
- Bateman, D. R., & Zidonis, F. J. (1964). The effect of a knowledge of generative grammar upon the growth of language complexity. Columbus: The Ohio State University.
- Beach, R. (1976). Self-evaluation strategies of extensive revisers and nonrevisers. College Composition and Communication, 27, 160-164.
- Beach, R. (1979). The effects of between-draft teacher evaluation versus student self-evaluation on high school students' revising of rough drafts. Research in the Teaching of English, 13, 111-119.
- Beach, R., & Bridwell, L. (Eds.). (1984). New directions in composition research. New York: Guilford Press.
- Becker, S. L., & Dallinger, C. A. (1960). The effect of instructional methods upon achievements and attitudes in communication skills. Speech Monographs, 27, 70-76.
- Belanger, J. (1978). Reading skill as an influence on writing skill. Unpublished doctoral dissertation, University of Alberta, Canada.
- Benedict, R. (1934). Patterns of culture. New York: Houghton Mifflin.
- Bereiter, C. (1960). Verbal and ideational fluency in able 10th graders. Journal of Educational Psychology, 51, 337-345.
- Bereiter, C., & Scardamalia, M. (1978). From conversation to composition. In R. Glaser (Ed.), Advances in instructional psychology (pp. 10-12). Hillsdale, NJ: Erlbaum.

- Bereiter, C., & Scardamalia, M. (1983). Levels of inquiry in writing research. In P. Mosenthal, L. Tamor, & S. A. Walmsley (Eds.), Research on writing: Principles and methods (pp. 3-25). London: Longman.
- Bereiter, C., Scardamalia, M., Anderson, V., & Smart, D. (1980). An experiment in teaching abstract planning in writing. Paper presented at the meeting of American Educational Research Association, Boston, MA.
- Berkenkotter, C. (1981). Understanding a writer's awareness of audience. College Composition and Communication, 32, 388-399.
- Berlin, J. (1987). Rhetoric and reality: Writing instruction in American colleges, 1900-1985. Carbondale: Southern Illinois University Press.
- Bettelheim, B. (1960). The informed heart: Autonomy in a mass age. Glencoe, IL: The Free Press.
- Bippus, A. C. (1977). The relationship of the quality of students' written language, productivity of writing, and reading comprehension in grades four and six (Doctoral dissertation, University of Virginia, 1977). Dissertation Abstracts International. 38, 3993A.
- Birnbaum, J. C. (1981). A study of reading and writing behaviors of selected fourth grade and seventh grade students (Doctoral dissertation, Rutgers University, 1981). Dissertation Abstracts International, 42, 152A.
- Bissex, G. (1980). Gyns at work: A child learns to read and write. Cambridge, MA: Harvard University Press.
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). Handbook of formative and summative evaluation of student learning. New York: McGraw-Hill.
- Blumer, H. (1969). Symbolic interactionism. Englewood Cliffs, NJ: Prentice-Hall.
- Boder, D. P. (1940). The adjective-verb quotient: A contribution to the psychology of language. Psychological Record, 3, 310-343.
- Bogdan, R. C., & Biklen, S. K. (1982). Qualitative research for education: An introduction to theory and methods. Boston: Allyn and Bacon.

- Bossone, R., & Quitman, L. A. (1976). A strategy for coping with high school and college remedial English problems. (ERIC Document Reproduction Service No. ED 130 268)
- Bracewell, R. J., Scardamalia, M., & Bereiter, C. (1978). The development of audience awareness in writing. Paper presented at the annual meeting of American Educational Research Association, Toronto, Canada.
- Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. American Educational Research Journal, 5, 437-474.
- Braddock, R., Lloyd-Jones, R., & Schoer, L. (1963). Research in written composition. Urbana, IL: National Council of Teachers of English.
- Bredo, E., & Feinberg, W. (Eds.). (1982). Knowledge and values in social and educational research. Philadelphia: Temple University Press.
- Bretzing, B., & Kulhavy, R. (1979). Notetaking and depth of processing. Contemporary Educational Psychology, 4, 145-153.
- Bridge, C., Hiebert, E. H., & Chesky, J. (1983). Classroom writing practices. In J. A. Niles and L. A. Harris (Eds.), Searches for meaning in reading/language processing and instruction (pp. 238-242). Rochester, NY: National Reading Conference, Inc.
- Bridwell, L. S. (1980). Revising strategies in twelfth grade students' transactional writing. Research in the Teaching of English, 14, 197-222.
- Britton, J. (Ed.) (1967). Talking and writing: A handbook for English teachers. London: Methuen.
- Britton, J. (1982). Shaping at the point of utterance. In G. Pradl (Ed.), Prospect and retrospect: Selected essays of James Britton (pp. 139-145). London: Heinemann.
- Britton, J., Burgess, T., Martin, N., McLeod, A., & Rosen, H. (1975). The development of writing abilities (11-18). London: Macmillan Education.
- Brostoff, A. (1979). Good assignments lead to good writing. Social Education, 43, 184-86.
- Bruner, J. S. (1971). The relevance of education. New York: W. W. Norton.

- Bryk, A. S., & Raudenbush, S. W. (1992). Hierarchical linear models: Applications and data-analysis methods. Newbury Park, CA: Sage.
- Buckley, M. H., & Boyle, O. (1983). Mapping and composing. In M. Myers and J. Grey (Eds.), Theory and practice in the teaching of composition (pp. 59-66). Urbana, IL: National Council of Teachers of English.
- Burstein, L., Miller, M. D., & Linn, R. L. (1979). The use of with-in group slopes as indices of group outcomes (CSE Report Series). Los Angeles: Center for the Study of Evaluation, University of California.
- Burton, D. L. (1973). Research in the teaching of English: The troubled dream. Research in the Teaching of English, 7, 160-189.
- Butterfield, H. (1960). The origins of modern science. New York: The Macmillan Company.
- Buxton, E. W. (1959). An experiment to test the effects of writing frequency and guided practice upon students' skill in written expression (Doctoral dissertation, Stanford University, 1958.) Dissertation Abstracts, 19, 709.
- Calhoun, J. (1971). The effect of analysis of essays in college composition classes on reading and writing skills (Doctoral dissertation, Boston University, 1971). Dissertation Abstracts International, 32, 1971A.
- Camp, G. (Ed.). (1982). Teaching writing: Essays from the Bay Area Writing Project. Berkeley, CA: Boynton-Cook.
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Chicago: Rand McNally College Publishing Co.
- Campbell, D. T., & Stanley, J. C. (1971). Experimental and quasi-experimental designs for research. Chicago: Rand McNally College Publishing Co.
- Campbell, M. L. (1976). An investigation of the relationship between secondary generative and receptive communication skills at the college freshman level (Doctoral dissertation, University of Southern Mississippi, 1976). Dissertation Abstracts International, 37, 5655A.
- Carrell, P. L., & Eisterhold, J. C. (1983). Schema theory and ESL reading pedagogy. TESOL Quarterly, 17, 553-573.

- Carroll, J. B. (1961, February). Neglected areas in educational research. Paper presented at a meeting of the American educational Research Association, Chicago.
- Carroll, J. M. (1984). Process into product: Teacher awareness of how the writing process affects students' written products. In R. Beach, & L. Bridwell (Eds.), New directions in composition research (pp. 315-333). New York: Guilford Press.
- Cast, B. M. D. (1940). The efficiency of different methods of marking English composition. British Journal of Educational Psychology, 9, 257-269.
- Chen, J. (1992). An experimental study of whole language instruction and basal reader instruction for the junior high English as a second language (ESL) students in Taiwan, Republic of China (Doctoral dissertation, University of Northern Colorado, 1991). Dissertation Abstracts International, 52, 2817A.
- Chesky, J., & Hiebert, E. H. (1987). The effects of prior knowledge and audience on high school students' writing. The Journal of Educational Research, 80, 304-313.
- Chotlos, J. W. (1944). A statistical and comparative analysis of individual written language samples. Doctoral dissertation, University of Iowa, 1942, pp. 14-19. Published in Psychological Monographs, 56, 77-110.
- Christensen, F. (1967). Notes toward a new rhetoric: Six essays for teachers. New York: Harper & Row.
- Christiansen, M. A. (1965). Tripling writing and omitting readings in freshman English: An experiment. College Composition and Communication, 16, 122-124.
- Clark, J. D. (1935). A four-year study of freshman English. English Journal (College edition), 24, 404-410.
- Cochran, W. G. (1937). Problems arising in the analysis of a series of similar experiments. Journal of the Royal Statistical Society, Supplement 4: 102-118.
- Cochran, W. G. (1963). Sampling techniques. New York: John Wiley & sons, Inc.
- Cohen, J. (1977). Statistical power analysis for the behavioral sciences (rev. ed.). Orlando, FL: Academic Press.

- Cole, M., & Scribner, S. (1974). Culture and thought: A psychological introduction. New York: Scribner's.
- Collins, C. (1979). The effect of writing experiences in the expressive mode upon the reading, self-esteem, attitudes, and academic achievement of freshmen in a college reading course (Doctoral dissertation, Rutgers University, 1979). Dissertation Abstracts International, 40, 3804A.
- Combs, W. (1979). Examining the fit of practice in syntactic manipulation and scores of reading comprehension. In D. Daiker, A. Kerek, & M. Morenberg (Eds.), Sentence-combining and the teaching of English (pp. 50-55). Conway, AR: University of Central Arkansas.
- Connolly, F. X. (1953). A rhetoric casebook. New York: Harcourt Brace Jovanovich.
- Connors, R. A. (1981). The rise and fall of the modes of discourse. College Composition and Communication, 32, 444-455.
- Cooper, C. R. (1973). Students learn to read fiction by writing it. California English Journal, 9, 7-17.
- Cooper, C. (1975). Measuring growth in writing. English Journal, 64(3), 111-120.
- Cooper, C., & Odell, L. (1977). Evaluating writing: Describing, meaning, judging. Urbana, IL: National Council of Teachers of English.
- Cooper, C., & Odell, L. (Eds.). (1978). Research on composing: Points of departure. Urbana, IL: National Council of Teachers of English.
- Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. Review of Educational Research, 52, 291-302.
- Cooper, M. (1984). The pragmatics of form: How do writers discover what to do then? In R. Beach, & L. Bridwell (Eds.), New directions in composition research (pp. 109-126). New York: Guilford Press.
- Corbett, E. P. J. (1965). Classical rhetoric for the modern student. New York: Oxford University Press.

- Crowhurst, M. (1981). Cohesion in argumentative prose written by sixth-, tenth-, and twelfth-graders. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, CA.
- Crowhurst, M. (1983a). Persuasive writing at grads 5, 7, and 11: A cognitive-developmental perspective. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.
- Crowhurst, M. (1983b). Revision strategies of students at three grade levels. (Final report). Educational Research Institute of British Columbia, Canada.
- Crowhurst, M., & Piche, G. L. (1979). Audience and mode of discourse effects on syntactic complexity in writing at two grade levels. Research in the Teaching of English, 18, 383-417.
- Culbertson, J. A., & Hencley, S. P. (Eds.). (1963). Educational research: New perspectives. Danville, IL: The Interstate Printers and Publishers, Inc.
- Daiker, D. A., Kerek, A., & Morenberg, M. (Eds.). (1979). Sentence-combining and the teaching of writing. Conway, AR: University of Central Arkansas.
- Daiker, D. A., Kerek, A., & Morenberg, M. (Eds.). (1985). Sentence-combining: A rhetorical perspective. Carbondale and Edwardsville, IL: Southern Illinois University Press.
- Daly, J., & Hexamer, A. (1983). Statistical power in research in English education. Research in the Teaching of English, 17, 156-164.
- Daly, J., & Miller, M. D. (1975). The empirical development of an instrument to measure writing apprehension. Research in the Teaching of English, 9, 272-289.
- Daly, J., & Wilson, D. A. (1983). Writing apprehension, self-esteem, and personality. Research in the Teaching of English, 17, 327-342.
- D'Angelo, J. (1977). Predicting reading achievement in a senior high school from intelligence, intelligence, and informative writing (Doctoral dissertation, University of Pittsburgh, 1977). Dissertation Abstracts International, 38, 2027A.

- Davidson, D. M. (1976). Development and validation of a diagnostic examination of ability to subordinate in writing for ESL college students (Doctoral dissertation, New York University, 1976). Dissertation Abstracts International, 37, 5790A.
- De Beaugrande, R. (1980). Text, discourse and process: Toward a multidisciplinary science of texts. Norwood, NJ: Ablex Publishing Corporation.
- De Beaugrande, R., & Dressler, W. (1981). Introduction to text linguistics. New York: Longman.
- Dempsey, A. V. (1949). An analysis of the errors made in the transcription of the shorthand. (Doctoral dissertation, University of Northern Colorado, 1949). Dissertations of Colorado State College, 11, 29.
- Dessner, L. E. (1991). English as a second language college writers' revision responses to teacher-written comments (Doctoral dissertation, University of Pennsylvania, 1991). Dissertation Abstracts International, 52, 827A.
- DeVries, T. (1970). Reading, writing frequency, and expository writing. Reading Improvement, 7, 14-15, 19.
- Diederich, P. (1957). The problem of grading essays. Princeton, New Jersey: Educational Testing Service.
- Diederich, P. (1965). Grading and measuring. In A. Jewett (Ed.), Improving English composition (pp. 81-91). Washington, DC: National Educational Association.
- Diederich, P. (1974). Measuring growth in English. Urbana, IL: National Council of Teachers of English.
- Dillon, G. L. (1981). Constructing texts: Elements of a theory of composition and style. Bloomington, IN: Indiana University Press.
- DiStefano, P., & Howie, S. (1979). Sentence weights: An alternative to the T-unit. English Education, 11, 98-101.
- Dixon, W. J., & Massey, F. K. (1969). Introduction to statistical analysis, New York: McGraw-Hill.
- Dobbert, M. L. (1982). Ethnographic research: Theory and application for modern schools and societies. New York: Praeger.

- Doctorow, M., Wittrock, M. C., & Marks, C. (1978). Generative processes in reading comprehension. Journal of Educational Psychology, 70, 109-118.
- Dolch, E. W. (1956). School research in reading. Elementary English, 33, 76-80.
- Donelson, K. L. (1967). Variables distinguishing between effective and ineffective writers at the tenth grade. Journal of Experimental Education, 35, 37-41.
- Donlan, D. (1974). Teaching writing in the content areas: Eleven hypotheses from a teacher survey. Research in the Teaching of English, 8, 250-262.
- Dumas, E. (1942). Syllabic errors in certain misspelled words (Doctoral dissertation, Colorado State College of Education, 1942). Dissertations of Colorado State College, 4, 28.
- Dynes, J. J. (1932). Comparison of two methods of studying history. Journal of Experimental Education, 1, 42-45.
- Dyson, A. H. (1982). Reading, writing, and language: Young children solving the written language puzzle. Language Arts, 59, 829-838.
- Dziombak, C. E. (1991). Searching for collaboration in the ESL computer lab and the ESL classroom (Doctoral dissertation, 1990). Dissertation Abstracts International, 51, 2296A.
- Ebel, R. L. (1951). Estimation of the reliability of ratings. Psychometrika, 16, 407-424.
- Egler, F. E. (1942). Vegetation as an object of study. Philosophy of Science, 9, 245-260.
- Elbaz, F. (1981). The teacher's 'practical knowledge': Report of a case study. Curriculum Inquiry, 11, 43-71.
- Elley, W. V., Barham, I. H., Lamb, H., & Wyllie, M. (1976). The role of grammar in a secondary school English curriculum. Research in the Teaching of English, 10, 5-21.
- Emig, J. (1971). The composing process of twelfth graders (Report No. 13). Urbana, IL: National Council of Teachers of English.

- Emig, J. (1977). Writing as a mode of learning. College Composition and Communication, 28, 122-127.
- Emig, J. (1982). Inquiry paradigms and writing. College Composition and Communication, 33, 64-75.
- Emig, J. (1983). The web of meaning: Essays on writing, teaching, learning, and thinking. Upper Montclair, NJ: Boynton/Cook.
- Eurich, A. C. (1931). The reading abilities of the college students: An experimental study. In College problems series. Minneapolis: University of Minnesota Press.
- Evanechko, P., Ollila, L., & Armstrong, R. (1974). An investigation of the relationship between children's performance in written language and their reading ability. Research in the Teaching of English, 8, 315-326.
- Evans, J. W. (1939). The social importance and the pupil control of certain punctuation variants (Doctoral dissertation, University of Iowa, 1939). Doctoral Dissertations Accepted by American Universities, 6, 72.
- Evans, R. (1979). The relationship between the reading and writing of syntactic structures. Research in the Teaching of English, 13, 129-136.
- Ezer, H. (1990). Writing development in a first grader: A case study (Doctoral dissertation, Boston University, 1990). Dissertation Abstracts International, 51, 1143A.
- Fahnestock, J. (1983). Semantic and lexical coherence. College Composition and Communication, 34, 400-416.
- Faigley, L., Daly, J., & Witte, S. (1981). The effects of writing apprehension on writing performance and competence. Journal of Educational Research, 75, 16-21.
- Feibleman, J. K. (1954). Theory of integrative levels. British Journal of Philosophy of Science, 5, 59-66.
- Felland, N. (1980). A national study of the level of composition achievement (superior/average) of twelfth grade composition students and selected personal characteristics/environmental factors (Doctoral dissertation, Northern Illinois University, 1980). Dissertation Abstracts International, 41, 3037A.
- Ferguson, G. A. (1981). Statistical analysis in psychology and education (5th ed.). New York: McGraw-Hill.

- Ferris, D. R. (1992). Syntactic and lexical characteristics of ESL student writing: A multidimensional study (Doctoral dissertation, University of Southern California, 1991). Dissertation Abstracts International, 52, 2907A.
- Fishco, D. T. (1966). A study of the relationship between creativity in writing and comprehension in reading of selected seventh grade students (Doctoral dissertation, Lehigh University, 1966). Dissertation Abstracts International, 27, 3220A.
- Fisher, R. A. (1932). Statistical methods for research workers (4th ed.). London: Oliver and Boyd.
- Flower, L. (1979). Writer-based prose: A cognitive basis for problems in writing. College English, 41, 19-37.
- Flower, L., & Hayes, J. (1977). Problem solving strategies and the writing process. College English, 39, 449-461.
- Flower, L., & Hayes, J. (1980a). The dynamics of composing: Making plans and juggling constraints. In L. W. Gregg & E. R. Steinberg (Eds.), Cognitive processes of writing (pp. 31-50). Hillsdale, NJ: Lawrence Erlbaum.
- Flower, L., & Hayes, J. (1980b). Cognition of discovery: Defining a rhetorical problem. College Composition and Communication, 31, 21-32.
- Foley, J. J. (1971). Evaluation of learning in writing. In B. S. Bloom, et al., Handbook on formative and summative evaluation of student learning (pp. 767-813). New York: McGraw-Hill.
- Foster, A. N. (1943). The test-study method versus the study-test method of teaching spelling (Doctoral dissertation, Colorado State College of Education, 1943). Dissertations of Colorado State College, 5, 22.
- Freud, S. (1956). The relation of the poet to day-dreaming. In I. F. G. Duff (Trans.), Delusion and dream and other essays by Sigmund Freud (pp. 122-133). Boston, MA: Beacon Press.
- Fries, C. (1945). Teaching and learning English as a second language. Ann Arbor: University of Michigan Press

- Fuller, K. M. (1974). An investigation of the relationship between reading achievement and oral and written language of students enrolled in reading and English classes at Gainesville Junior College (Doctoral dissertation, University of Georgia, 1974). Dissertation Abstracts International, 35, 6692A.
- Fulwiler, T. E. (1984). How well does writing across the curriculum work? College English, 46, 113-126.
- Gay, L. R. (1987). Educational research: Competencies for analysis and application (3rd ed.). Columbus, OH: Merrill Publishing company.
- Giffen-Glad, C. E. (1991). Homework performance patterns among Hispanic secondary students (Doctoral dissertation, University of La Verne, 1991). Dissertation Abstracts International, 52, 2817A.
- Gillie, P. J. (1957). A simplified formula for measuring abstraction in writing. Journal of Applied Psychology, 41, 214-217.
- Gioconia, R., & Hedges, L. V. (1982). Identifying features of effective open education. Review of Educational Research, 52, 579-602.
- Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine.
- Glass, G. (1976). Primary, secondary, and meta-analysis of research. Educational Researcher, 5, 3-8.
- Glass, G. (1977). Synthesizing empirical research: Meta-analysis. Review of Research in Education, 47, 351-379.
- Glass, G. (1978). Integrating findings: The meta-analysis of research. In L. S. Shulman (Ed.), Review of research in education (5th Vol.), (pp. 351-379). Itasca, IL: F. E. Peacock.
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage.
- Glock, C. Y. (1967). Survey research in the social sciences. New York: Russell Sage Foundation.

- Glover, J., Plake, B., Roberts, B., Zimmer, J., & Palmere, M. (1981). Distinctiveness of encoding: The effects of paraphrasing and drawing inferences on memory from prose. Journal of Educational Psychology, 73, 736-744.
- Goetz, J. P., & LeCompte, M. D. (1984). Ethnography and qualitative design in educational research. Orlando, FL: Academic Press, Inc.
- Godshalk, F., Swineford, F., & Coffman, W. (1966). The measurement of writing ability. New York: College Entrance Examination Board.
- Goodman, K. (1967). Reading: A psycholinguistic guessing game. Journal of the Reading Specialist, 6, 126-135.
- Graves, D. H. (1975). An examination of the writing processes of seven-year-old children. Research in the Teaching of English, 9, 227-241.
- Graves, D. H. (1980). Research update: A new look at writing research. Language Arts, 57, 913-919.
- Graves, D. H. (1983). Writing: Teachers and children at work. New Hampshire: Heinemann Educational Books.
- Graves, D. H., & Murray, D. M. (1980). Revision: In the writer's workshop and in the classroom. Journal of Education, 162(2), 38-56.
- Green, B., & Hall, J. (1984). Quantitative methods for literature review. Annual Review of Psychology, 35, 37-53.
- Gregg, L. W., & Steinberg, E. R. (Eds.). (1980). Cognitive processes of writing. Hillsdale, NJ: Lawrence Erlbaum.
- Grimmer, F. (1970). The effects of an experimental program in written composition on the writing of second-grade children (Doctoral dissertation, University of Georgia, 1970). Dissertation Abstracts International, 31, 5666A.
- Grobe, C. (1981). Syntactic maturity, mechanics, and vocabulary as predictors of quality ratings. Research in the Teaching of English, 15, 75-85.
- Grobe, S. F., & Grobe, C. H. (1977). Reading skills as a correlate of writing ability in college freshmen. Reading World, 16, 50-54.

- Guba, E. G. (1963). Experiments, studies, surveys, and investigations. In J. A. Culbertson & S. P. Hencley (Eds.), Educational research: New perspectives (pp. 237-248). Danville, IL: The Interstate Printers and Publishers, Inc.
- Haber, S. B. (1992). Multiple perspectives of group interaction in an ESL writing class (Doctoral dissertation, Columbia University Teachers College, 1991). Dissertation Abstracts International, 52, 2443A.
- Hafernik, J. J. (1991). Relationships among English writing experience, contrastive rhetoric, and English expository prose of L(1) and L(2) college writers (Doctoral dissertation, University of San Francisco, 1990). Dissertation Abstracts International, 51, 4007A.
- Hall, E. T. (1959). The silent language. Garden City, NY: Doubleday.
- Hall, E. T. (1966). The hidden dimension. Garden City, NY: Doubleday.
- Halliday, M. A. K., & Hasan, R. (1976). Cohesion in English. London: Longman.
- Halliday, M. A. K., McIntosh, A., & Stevens, P. (1964). The linguistic sciences and language teaching. Bloomington: Indiana University Press.
- Hansen, M. H., Hurwitz, W. N., & Madow, W. G. (1953). Sample survey methods and theory. New York: John Wiley & Sons, Inc.
- Harris, R. J. (1962). An experimental inquiry into the functions and values of formal grammar in the teaching of English, with special reference to the teaching of correct written English to children aged twelve to fourteen. Unpublished doctoral dissertation, University of London, Britain.
- Hart, J. S. (1855). The Anglo-Saxon element in the English language. The American Journal of Education, 1(1), 33-60.
- Hartog, P., & Rhodes, E. C. (1936). The marks of examiners. London: Macmillan Company.
- Haskins, J. B. (1960). Validation of the abstraction index as a tool for content-effects analysis and content analysis. Journal of Applied Psychology, 44, 102-106.

- Hatch, G. (1991, March). Reviving the rodential model for composition: Robert Zoellner's alternative to Flower and Hayes. Paper presented at the meeting of the Conference on College Composition and Communication, Boston, MA.
- Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), Cognitive processes in writing (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Hays, E. (1936). College entrance requirements in English: Their effects on school. New York: Teachers College, Columbia University.
- Hays, J. N., Roth, P. A., Ramsey, J. R., & Foulke, R. D. (Eds.). (1983). The writer's mind: Writing as a mode of thinking. Urbana, IL: National Council of Teachers of English.
- Hayward, N. M. (1991). The reluctant writer: A descriptive study of student behavior and motivation in the composition class (Doctoral dissertation, Indiana University of Pennsylvania, 1991). Dissertation Abstracts International, 52, 2054A.
- Healy, M. A. K. (1981). Purpose in learning to write: An approach to writing in three curricular areas. In C. H. Fredericksen and J. F. Dominie (Eds.), Writing: the nature, development, and teaching of written communication: Vol. 2. Processes, development and communication (pp. 223-233). Hillsdale, NJ: Lawrence Erlbaum.
- Heaphey, J. (1963). The problem of theory and case research. In J. A. Culbertson & S. P. Hencley (Eds.) Educational research: New perspectives (pp. 265-276). Danville, IL: The Interstate Printers & Publishers, Inc.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. Journal of Educational Statistics, 6, 107-128.
- Hedges, L. V. (1982a). Estimating effect size from a series of independent experiments. Psychological Bulletin, 92, 490-499.
- Hedges, L. V. (1982b). Fitting categorical models to effect sizes from a series of experiments. Journal of Educational Statistics, 7, 119-137.

- Heil, H. (1977). The development of selected language variables in two modes of writing and their relationship to reading comprehension in the primary grades (Doctoral dissertation, Hofstra University, 1976). Dissertation Abstracts International, 38, 137A.
- Heller, M. (1980). The reading-writing connection: An analysis of the written language of university freshmen at two reading levels (Doctoral dissertation, Oklahoma State University, 1979). Dissertation Abstracts International, 40, 4452A.
- Heys, F. (1962). The theme-a-week assumption: A report on an experiment. English Journal, 51, 320-322.
- Hilgers, T. L. (1982). Experimental control and the writing stimulus: The problem of unequal familiarity with content. Research in the Teaching of English, 16, 381-190.
- Hillocks, G., Jr. (1981). The response of college freshmen to three modes of instruction. American Journal of Education, 89, 373-395.
- Hillocks, G., Jr. (Ed.). (1982). The English curriculum under fire. Urbana, IL: National Council of Teachers of English.
- Hillocks, G., Jr. (1984). What works in teaching composition: A meta-analysis of experimental treatment studies. American Journal of Education, 93, 133-170.
- Hillocks, G., Jr. (1986). Research on written composition: New directions for teaching. Urbana, IL: ERIC Clearinghouse on Reading and communication Skills and National Conference on Research in English.
- Hoetker, J. (1982). Effects of the essay topics on student writing: Review of the literature. Tallahassee: Florida State University.
- Holbrook, D. (1961). English for the rejected: Training literacy in the lower streams of the secondary school. Cambridge, England: Cambridge University Press.
- Hou, Z. (1987, January). English teaching in China: Problems and perspectives. TESOL News, pp. 25-27.
- Huber, P. J. (1977). Robust statistical procedures. Philadelphia: Society for Industrial and Applied Mathematics.

- Huddleson, E. M. (1954). Measurement of writing ability at the college-entrance level: Objective vs. subjective testing techniques. Journal of Experimental Education, 22, 165-213.
- Hunt, K. (1965). Grammatical structures written at three grade levels. Urbana, IL: National Council of Teachers of English.
- Hunt, K. (1970). Syntactic maturity in school children and adults. Monographs of the Society for Research in Child Development, 35(Serial No. 134).
- Hunt, K. (1976). A synopsis of clause-to-sentence length factors. In R. Graves (Ed.), Rhetoric and composition: A sourcebook for teachers (pp. 110-117). Rochelle Park, NJ: Hayden.
- Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). Meta-analysis: Cumulating research findings across studies. Beverly Hills, CA: Sage.
- Hwang, H. J. (1992). The development of English writing proficiency among Korean students from ninth-grade to college juniors (Doctoral dissertation, University of California, Santa Barbara, 1991). Dissertation Abstracts International, 52, 3814A.
- Hyman, H. H. (1955). Survey design and analysis. New York: The Free Press.
- Iker, H. P., & Harway, N. (1965). A computer approach towards the analysis of content. Behavioral Science, 10, 173-182.
- Jackson, G. (1980). Methods for integrative reviews. Review of Educational Research, 50, 438-460.
- Jacobson, L. (1969). Rhetoric and stylistics: Some basic issues in the analysis of discourse. College Composition and Communication, 20, 314-328.
- Jencke, G. (1935). A study of précis writing as a composition technique. Contribution to education (Report No. 644). New York: Bureau of Publications, Teachers College, Columbia University.
- Jensen, J. M. (Ed.). (1984). Composing and comprehending. Urbana, IL: National Council of Teachers of English.

- Johnson, N. (1981). A comparison of syntactic writing maturity with reading achievement (Doctoral dissertation, East Texas State University, 1980). Dissertation Abstracts International, 41, 4346A.
- Jordan, C. A. (1992). Measurement of giftedness of elementary school children and identification of black or Hispanic educationally disadvantaged gifted children (Doctoral dissertation, Texas Woman's University, 1991). Dissertation Abstracts International, 52, 2817A.
- Judd, E. L. (1978). Language policy, curriculum development, and TESOL instruction: A search for compatibility. In M. H. Long, and J. C. Richards (Eds.), Methodology in TESOL: A book of readings (pp. 3-9). New York, NY: Newbury House Publishers.
- Judy, S. N. (1967). The teaching of English composition in American secondary schools, 1850-1893 (Doctoral dissertation, Northwestern University, 1967). Dissertation Abstracts, 28, 2210A.
- Kaldegg, G. (1950). Substance symbolism: A study in language psychology. Journal of Experimental Education, 28, 331-342.
- Kerlinger, F. N. (1973). Foundations of behavioral research. New York: Holt, Reinhart, and Winston.
- Kim, W. (1991). The role of metacognitive skills in young ESL students' writing revisions (Doctoral dissertation, University of Hawaii, 1991). Dissertation Abstracts International, 52, 1668A.
- Kimoto, B. (1951). A quantitative analysis of grammatical categories as a measure of dominance (Doctoral dissertation, Boston University, 1951). Doctoral Dissertations Accepted by American Universities, 18, 125.
- Kincaid, G. L. (1953). Some factors affecting variations in the quality of students' writing (Doctoral dissertation, Michigan State College, 1953). Dissertation Abstracts, 13, 733,
- Kineavy, J. E. (1969). The basic aims of discourse. College Composition and Communication, 20, 297-304.
- Kineavy, J. E. (1971). A theory of discourse. Englewood Cliffs, NJ: Prentice-Hall.

- King, M., & Rentel, V. (1979). Towards a theory of early writing development. Research in the Teaching of English, 13, 243-253.
- Kish, L. (1965). Survey sampling. New York: John Wiley & Sons, Inc.
- Kitchin, D. A. (1991) Case study of ESL community college students using computer-based writing tools in a composition course (Doctoral dissertation, University of San Francisco, 1991). Dissertation Abstracts International, 52, 1240A.
- Kitzhaber, A. R. (1953). Rhetoric in American colleges, 1850-1900 (Doctoral dissertation, University of Washington, 1953). Dissertaion Abstracts, 13, 810.
- Kitzhaber, A. R. (1963). Themes, theories, and therapy: The teaching of writing in college. New York: McGraw-Hill.
- Kohl, H. (1967). 36 children. New York: New American Library.
- Krashen, S. (1981). Second language acquisition and learning. Oxford: Pergamon.
- Kuhn, T. S. (1962). The structure of scientific revolution. Chicago: The University of Chicago Press.
- Kulhavy, R., Dyer, J., & Silver, L. (1975). The effects of notetaking and test expectancy on the learning of text material. Journal of Educational Research, 68, 363-365.
- Kulik, J. (1983). Book review. [Review of Meta-analysis in social research]. Evaluation News, 4, 101-105.
- LaBrant, L. L. (1933). A study of certain language developments of children in grades four to twelve, inclusive. Genetic Psychology Monographs, 14, 387-491.
- LaCampagne, R. J. (1968). A national study of selected attitudes and approaches to writing of twelfth-grade students with superior writing performance versus those with average writing performance (Doctoral dissertation, University of Illinois, 1968). Dissertation Abstracts International, 30, 212.

- Lam, C. Y. P. (1992). Revision processes of college ESL students: How teacher comments, discourse types, and writing tools shape revision (Doctoral dissertation, University of Georgia, 1991). Dissertation Abstracts International, 52, 4248A.
- Langer, J. A. (1984). The effects of available information on responses to school writing tasks. Research in the Teaching of English, 18, 27-44.
- Langer, J., & Applebee, A. (1982). Writing and learning in the secondary school, Proposal to the National Institute of Education. Stanford, CA: Stanford University.
- Langer, J., & Applebee, A. (1986). Writing and learning in the secondary school, Final Report to the National Institute of Education. Stanford, CA: Stanford University.
- Lauer, J. M., & Asher, J. W. (1988). Composition research: Empirical designs. New York: Oxford University Press.
- Lazarsfeld, P. F., & Rosenberg, M. (Eds.). (1955). The language of social research. New York: The Free Press.
- Lazdowski, W. (1976). Determining reading grade levels from analysis of written compositions (Doctoral dissertation, New Mexico State University, 1976). Dissertation Abstracts International, 37, 1504A.
- Lee, C. (1991). Korean high school seniors' oral and literate comprehension and production skills in English (Doctoral dissertation, Harvard University, 1990). Dissertation Abstracts International, 51, 3004A.
- Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. Sociology of Education, 62, 172-192.
- Leonard, S. A. (1932). Current English usage. Chicago: Inland Press.
- Li, P. (1992). Situational simulation of teaching English methodology (Doctoral dissertation, Drake University, 1991). Dissertation Abstracts International, 53, 1038A.
- Light, R. J., & Pillemer, D. B. (1984). Summing up: The science of reviewing research. Cambridge, MA: Harvard University Press.

- Likert, R. (1932). A technique for the measurement of attitudes. Archives of Psychology (No. 140).
- Liu, X. (1988, December). An analysis of factors that hinder innovations in China. TESOL Reporter, pp. 70-77.
- Loban, W. (1963). The language of elementary children (Report No. 1). Urbana, IL: National Council of Teachers of English.
- Loban, W. (1976). Language development: Kindergarten through grade twelve. Urbana, IL: National Council of Teachers of English.
- Lockheed, M. E., & Longford, N. T. (1989). A multilevel model of school effectiveness in a developing country. Washington, DC: World Bank.
- Lofland, J. (1976). Doing social life: The qualitative study of human interaction in natural settings. New York: Wiley.
- Lopez, C. L. (1990). The relationships between practical value of English and instrumental motivation among Puerto Rican ESL students (Doctoral dissertation, Fordham University, 1990). Dissertation Abstracts International, 51, 1939A.
- Luria, A. R., & Yudovich, F. I. (1971). Speech and the development of mental processes in the child. Baltimore: Penguin.
- Lutkus, A. (1987). Problems in measuring syntactic development: T-units vs. sentence weight. Journal of Teaching Writing, 6, 49-67.
- Lyman, R. L. (1922). English grammar in American schools before 1850. Chicago: University of Chicago Libraries.
- Maat, D. (1977). An enquiry into empirical relationships between the reading and writing of exposition and argument (Doctoral dissertation, State University of New York at Albany, 1977). Dissertation Abstracts International, 38, 4631-32A.
- Maloney, H. G. (1967). An identification of excellence in expository composition performance in a selected 9A population with an analysis of reasons for superior performances (Doctoral dissertation, Columbia University, 1967). Dissertation Abstracts International, 28, 3564A.

- Manis, J., & Meltzer, B. (Eds.). (1967). Symbolic interaction: A reader in social psychology. Boston, MA: Allyn and Bacon.
- Markels, R. B. (1983). Cohesion paradigms in paragraphs. College English, 45, 450-464.
- Martin, A. R. (1991). Teaching writing as a tool for learning with adult ESL students: A case study (Doctoral dissertation, University of Hawaii, 1990). Dissertation Abstracts International, 51, 3304A.
- Martin, H., Ohmann, R., & Wheatley, J. (1969). The logic and rhetoric of exposition. New York: Holt, Rinehart and Winston.
- Martin, J. (1976). Writing and learning across the curriculum. London: Ward Lock Education.
- Matheson, H. N. (1960). A study of errors made in paragraphs written by grade 12 students on the June, 1953, English 40 (Language) University Entrance examination, British Columbia Department of Education. Unpublished master's thesis, University of British Columbia, Canada.
- Mathews, E., Larsen, R., & Butler, G. (1945). Experimental investigation of the relation between reading training and achievement in college composition classes. Journal of Educational Research, 38, 499-505.
- Matsuhashi, A. (1982). Explorations in the real-time production of written discourse. In M. Nystrand (Ed.), What writers know: The language, process, and structure of written discourse (pp. 269-290). New York: Academic Press.
- May, P. N. (1992). Field analysis of the integrated activity learning sequence approach to science instruction at the elementary school level. Part I: Student performance (Doctoral dissertation, Temple University, 1991). Dissertation Abstracts International, 52, 2405A.
- Mazzuca, S. A., Cohen, S. J., & Clark, C. M., Jr. (1981). Evaluating clinical knowledge across years of medical training. Journal of Medical Education, 56(2), 83-90.
- McBride, O. (1941). The teaching of English in the southern ante-bellum academy. Nashville, TN: George Peabody College for Teachers.

- McCann, T. M. (1989). Student argumentative writing knowledge and ability at three grade levels. Research in the Teaching of English, 23, 62-76.
- McColly, W. (1970). What does educational research say about the judging of writing ability. Journal of Educational Research, 64, 144-156.
- McCrum, R., Cran, W., & MacNeil, R. (1986). The story of English. New York: Elisabeth Sifton Books-Viking.
- McCulley, G. A. (1985). Writing quality, coherence, and cohesion. Research in the Teaching of English, 19, 269-282.
- Meckel, H. C. (1963). Research on teaching composition and literature. In N. L. Gage (Ed.), Handbook of research on teaching (pp. 966-1006). Chicago: Rand McNally and Co.
- Mellon, J. C. (1969). Transformational sentence-combining: A method for enhancing the development of syntactic fluency in English composition (Report No. 10). Urbana, IL: National Council of Teachers of English.
- Mellon, J. C. (1975). National assessment and the teaching of English. Urbana, IL: National Council of Teachers of English.
- Mesavoric, M. D., & Macko, D. (1969). Foundations for a scientific theory of hierarchical systems. In L. L. Whyte, A. G. Wilson, & D. Wilson (Eds.), Hierarchical structure (pp. 29-50). New York: Elsevier.
- Mesavoric, M. D., Macko, D., & Takahara, Y. (1970). Theory of hierarchical multilevel systems. New York: Academic Press.
- Meyer, B. (1975). The organization of prose and its effect on memory. Amsterdam: North Holland.
- Meyer, B. (1982). Reading research and the composition teacher: The importance of plans. College Composition and Communication, 33, 37-49.
- Meyers, M. (Ed). (1985). The teacher-researcher: How to study writing in the classroom. Urbana, IL: National Council of Teachers of English.

- Miller, J. W. (1958). An analysis of freshman writing at the beginning and end of a year's work in composition (Doctoral dissertation, University of Minnesota, 1958). Dissertation Abstracts, 19, 2029.
- Miller, M. (1974). A comparative study of two approaches to teaching freshmen remedial composition in a comprehensive community college (Doctoral dissertation, University of Kansas, 1974). Dissertation Abstracts International, 35, 7083.
- Mills, E. (1974). Children's literature and teaching written composition. Elementary English, 51, 971-973.
- Mischel, T. (1974). A case study of a twelfth-grade writer. Research in the Teaching of English, 8, 303-314.
- Moffet, J. (1968a). Teaching the universe of discourse. Boston, MA: Houghton Mifflin.
- Moffet, J. W. (1968b). A student-centered language arts curriculum, grades K-13: A handbook for teachers. New York: Houghton Mifflin.
- Monk, R. (1958). A study to determine the relationship between children's home environments and their school achievement in written English (Doctoral dissertation, University of Washington, 1958). Dissertation Abstracts, 19, 1619.
- Mooney, P. M. (1991). Variables affecting oral proficiency among Spanish-speaking adults learning English as a second language (Doctoral dissertation, State University of New York at Albany, 1991). Dissertation Abstracts International, 52, 1986A.
- Morgan, E. J. (1990). The Given-New Contract in the compositions of Malaysian ESL writers as an indicator of writing quality (Doctoral dissertation, State University of New York at Buffalo, 1990). Dissertation Abstracts International, 51, 1142A.
- Mullis, I. V. S., & Mellon, J. C. (1980). Guideline for describing three aspects of writing: Syntax, cohesion, and mechanics. Denver, CO: National Assessment of Educational Progress. (ERIC Document reproduction Service No. ED 145 482)
- Murray, D. M. (1968). A writer teaches writing: A practical method of teaching composition. Boston, MA: Houghton Mifflin.

- Mustapha, S. Z. (1992). Reading and writing in a specific environment: The Malaysian experience (Doctoral dissertation, Indiana University, 1991). Dissertation Abstracts International, 52, 3233A.
- Myers, M. (1980). A procedure for writing assessment and holistic scoring. Urbana, IL: National Council of Teachers of English.
- Nagle, J. (1972). The effects of a directed writing activity in eighth grade social studies instruction on general reading achievement and on social studies reading achievement (Doctoral dissertation, Temple University, 1972). Dissertation Abstracts International, 33, 1523A.
- National Assessment of Educational Progress. (1980). Procedural handbook: 1978-79 writing assessment (Report No. 10-4-40). Denver, CO: National Assessment of Educational Progress.
- National Commission of Education of the People's Republic of China. (1985). Guideline for Curricula of University English Majors. Beijing, China: National Commission of Education Printing Office.
- Newell, G. (1984). Learning from writing in two content areas: A case study/protocol analysis. Research in the Teaching of English, 18, 265-287.
- Newlun, C. (1930). Teaching children to summarize in fifth grade history. Contributions to education (Report No. 404). New York: Bureau of Publications, Teachers College, Columbia University.
- Ney, J. W. (1966). Applied linguistics in the seventh grade. English Journal, 55, 895-897.
- Nold, E. W., & Freedman, S. W. (1977). An analysis of readers' responses to essays. Research in the Teaching of English, 11(12), 164-174.
- Novikoff, A. B. (1945). The concept of integrative levels and biology. Science, 101, 209-215.
- Nunnally, J. C. (1967). Psychometric theory. New York: MacGraw-Hill.
- Nunnally, J. C. (1975). Psychometric theory 25 years ago and now. Educational Researcher, 4(10), 21-25.

- Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.
- Obenchain, A. (1971). Effectiveness of the precise essay question in programming the sequential development of written composition skills and the simultaneous development of critical reading skills. Unpublished master's thesis, George Washington University, Washington, DC.
- O'Donnel, R. C., Griffin, W. J., & Norris, R. C. (1967). Syntax of kindergarten and elementary school children: A transformational analysis (Report No. 8). Urbana, IL: National Council of Teachers of English..
- Oehlkers, W. (1971). The contributions of creative writing to reading achievement in the language experience approach (Doctoral dissertation, University of Iowa, 1971). Dissertation Abstracts International, 32, 6689A.
- O'Hare, F. (1971). Sentence-combining: Improving student writing without formal grammar instruction (Report No. 15). Urbana, IL: National Council of Teachers of English.
- Olson, D. (1977). From utterance to text. Harvard Educational Review, 47(3), 257-281.
- Olson, D., & Torrance, N. (1981). Learning to meet the requirement of the text: Language development in the schoolyears. In C. H. Fredericksen, & F. Dominic (Eds.), Writing: The nature, development and teaching of written communication: Vol. 2: Processes, development and communication (pp. 235-255). Hillsdale, NJ: Lawrence Erlbaum.
- Olson, J. K. (1980). Teacher constructs and curriculum change. Journal of Curriculum Studies, 12, 1-12.
- O'Neill, R. V., DeAngelis, D. L., Waide, J. B., & Allen, T. F. H. (1986). A hierarchical concept of ecosystems (Monographs in Population Biology, No. 23). Princeton, NJ: Princeton University Press.
- Ong, W. (1982). Orality and literacy: The technologizing of the word. London: Methuen.
- Oppenheim, A. N. (1966). Questionnaire design and attitude measurement. New York: Basic Books, Inc., Publishers.

- Ozmon, H. A., & Craver, S. M. (1981). Philosophical foundations of education (2nd ed.). Columbus, OH: Charles E. Merrill Publishing Company.
- Page, E. B. (1968). The use of the computer in analyzing student essays. International Review of Education, 14, 253-263.
- Page, E. B., & Paulus, D. H. (1968). The analysis of essays by computers. Storrs: University of Connecticut.
- Parker, R. P. (1979). From Sputnik to Dartmouth: Trends in the teaching of composition. English Journal, 68, 32-37.
- Pattie, H. H. (1973). (Ed.). Hierarchy theory. New York: Braziller.
- Patton, M. Q. (1980). Qualitative evaluation methods. Beverly Hills, CA: Sage.
- Pearson, K. (1933). On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. Biometrika, 25, 379-410.
- Pedersen, E. (1977). Improving syntactic and semantic fluency in the writing of language arts students through extended practice in sentence-combining (Doctoral dissertation, University of Minnesota, 1977). Dissertation Abstracts International, 38, 5892A.
- Perl, S. (1979). The composing processes of unskilled college writers. Research in the Teaching of English, 13, 317-336.
- Perron, J. (1977). The impact of mode on written syntactic complexity: Part IV-across-the-grades' differences and general summary. In Studies in Language Education (Report No. 30). Athens: Department of Language Education, University of Georgia.
- Pestalozzi, J. H. (1859). Selections from publications by Pestalozzi. The American Journal of Education, 7, 519-702.
- Piaget, J. (1955). The language and thought of the child (M. Gabain, Trans.). New York: World Publishing Co.
- Pianko, S. H. (1979). A description of the composing processes of college freshman writers. Research in the Teaching of English, 13, 5-22.

- Piexotto, H. (1946). The relationship of College Board Examination scores and reading scores for college freshman. Journal of Applied Psychology, 30, 406-411.
- Postman, N. (1967). Linguistics and the pursuit of relevance. English Journal, 56, 1160-1165.
- Purves, A. C. (1983). Language and processing: Reading and writing. College English, 45, 129-140.
- Purves, A. C., & Purves, W. (1986). Cultures, text models and the activity of writing: Speculations for future research. Urbana: Curricular Laboratory, University of Illinois at Urbana-Champaign.
- Purves, A. C., & Rippere, V. (1968). Elements of writing about a literary work: A study of response to literature. Urbana, IL: National Council of Teachers of English.
- Purves, A. C., & Tierney, R. (1983). On the nature of rhetorical communities. Urbana: Curriculum Laboratory, University of Illinois at Urbana-Champaign.
- Raudenbush, S. W., & Bryk, A. S. (1986). A hierarchical model for studying school effect. Sociology of Education, 59, 1-17.
- Raudenbush, S. W., & Willms, J. D. (1991). Schools, classrooms, and pupils: International studies of schooling from a multilevel perspective. San Diego, CA: Academic Press.
- Reed, C. (1971). Preschool children's knowledge of English phonology. Harvard Educational Review, 41, pp. 1-34.
- Reed, C. (1975). Children's categories of speech sounds in English. Urbana, IL: National Council of Teachers of English.
- Reed, W. M., Burton, J. K., & Kelly, P. P. (1985). The effects of writing ability and mode of discourse on cognitive capacity engagement. Research in the Teaching of English, 19, 283-297.
- Reed, W. M., Burton, J. K., & Vandett, N. M. (1988). Daly and Miller's writing apprehension test and Hunt's T-unit analysis: Two measurement precautions in writing research. Journal of Research and Development in Education, 21(2), 1-8.

- Reed, W. M., & Sherman, T. M. (1991). Using memory during prewriting: The comparative effects of external storage and idea organization on the syntactic complexity and quality of basic and honors students' narrative writing. Research and Teaching in Developmental Education, 7(2), 67-79.
- Reed, W. M., & Vandett, N. M. (1986). Initiation versus intensification: The effect of topic-type on the performance of basic writers. Morgantown: West Virginia University, College of Human Resources and Education. (ERIC Reproduction Service No. ED 263 611)
- Reed, W. M., Vandett, N. M., & Burton, J. K. (1984). The effects of writing skills, sex, and writing preparation on writing apprehension. FORUM in Reading and Language Education, 1, 45-60.
- Reichardt, C. S., & Cook, T. D. (1979). Beyond qualitative versus quantitative methods. In C. S. Reichardt & T. D. Cook (Eds.), Qualitative and quantitative methods in evaluation research (pp. 7-32). Beverly Hills, CA: Sage.
- Rivlin, H. N. (1938). The present status of research in functional grammar. English Journal, 27, 590-597.
- Robinson, H. M. (1946). Why pupils fail in reading: A study of causes and remedial treatment. Chicago: University of Chicago Press.
- Rockas, L. (1964). Modes of rhetoric. New York: St. Martin's Press.
- Rohman, D. G. (1965). Pre-writing: The stage of discovery in the writing process. College Composition and Communication, 16, 106-112.
- Roscoe, J. T. (1975). Fundamental research statistics for the behavioral sciences (2nd ed.). New York: Holt, Rinehart and Winston.
- Rosen, R. (1977). Complexity as a system property. General Systems, 3, 227-232.
- Rosenberg, M. (1968). The logic of survey analysis. New York: Basic Books, Inc., Publishers.
- Rosenthal, R. (1966). Experimenter effects in behavioral research. New York: Appleton-Century-Crofts.

- Rosenthal, R. (1984). Meta-analytic procedures for social research. Beverly Hill, CA: Sage.
- Rowe, J. S. (1961). The level-of-integration concept and ecology. Ecology, 42, 420-427.
- Rubin, D. L., & Piche, G. L. (1979). Development in syntactic and strategic aspects of audience adaptation skills in written persuasive communication. Research in the Teaching of English, 13, 293-316.
- Rumelhart D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, and W. E. Brewer (Eds.), Theoretical issues in reading comprehension (pp. 33-58). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rumelhart, D. E., & Ortony, A. (1977). The presentation of knowledge in memory. In R. C. Anderson, R. J. Spiro, and W. E. Montague (Eds.), Schooling and the acquisition of knowledge (pp. 99-135). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sager, C. (1973). Improving the quality of written composition through pupil use of rating scale (Doctoral dissertation, Boston University, 1973). Dissertation Abstracts International, 34, 1496A.
- Salisbury, R. (1934). A study of the effects of training in logical organization as a method of improving skills in study. Journal of Educational Research, 28, 241-254.
- Sanchez-Villamil, O. I. (1991). The effects of two metacognitive strategies on intermediate ESL college students' writing (Doctoral dissertation, Fordham University, 1991). Dissertation Abstracts International, 52, 1670A.
- Scardamalia, M. (1982). Assimilative processes in composition planning. Educational Psychologist, 17, 165-171.
- Scardamalia, M., & Bereiter, C. (1983). The development of evaluative, diagnostic, and remedial capabilities in children's composing. In M. Martlew (Ed.), The psychology of written language: A developmental approach (pp. 67-95). London: John Wiley & Sons.
- Scardamalia, M., & Bereiter, C. (1986). Written composition. In M. Wittrock (Ed.), Handbook of research on teaching (3rd ed.), (pp. 778-803). New York: Macmillan.

- Scheffé, H. (1953). A method for judging all contrasts in the analysis of variance. Biometrika, 40, 87-104.
- Schneider, V. (1971). A study of the effectiveness of emphasizing the teaching of reading skills to improve composition skills in remedial English classes at Kansas City Community Junior College (Doctoral dissertation, University of Kansas, 1971). Dissertation Abstracts International, 31, 6369A.
- Schonell, F. (1942). Backwardness in the basic subjects. Toronto: Clarke, Irwin.
- Scrafford, R. (1943). A study of the desirability of correlating English composition with social studies (Doctoral dissertation, Pennsylvania State College, 1942). Dissertations of the Pennsylvania State University, 6, 144.
- Shanahan, T. (1980). A canonical correlational analysis of learning to read and learning to write: An exploratory analysis (Doctoral dissertation, University of Delaware, 1980). Dissertation Abstract International, 41, 1007A.
- Siedow, M. (1973). Relationships between syntactic maturity in oral and written language and reading comprehension of materials of varying syntactic complexity (Doctoral dissertation, Indiana University, 1973). Dissertation Abstracts International, 34, 4890A.
- Silva, T. (1988). ESL composition instruction: Developments, issues, and directions. In B. Kroll (Ed.), Second language writing: Research insights for the classroom (pp. 11-23). New York: Cambridge.
- Simon, H. A. (1962). The architecture of complexity. Proceedings of American Philosophical Society, 106, 467-482.
- Simon, H. A. (1973). The organization of complex systems. In H. H. Pattee (Ed.), Hierarchy theory (pp. 3-27). New York: Brazziller.
- Singleton, C. M. (1961). Freedom to research. Elementary English, 38, 114-117, 121.
- Skinner, B. F. (1978). Reflections on behaviorism and society. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Small, R., Jr. (1979). The YA novel in the composition program, Part II. English Journal, 68(7), 75-77.

- Smith, D. V. (1931). Class size in high school English: Methods and results. Minneapolis: University of Minnesota Press.
- Smith, F. (1981). Writing and the writer. New York: Holt, Rinehart & Winston.
- Smith, L. M. (1979). An evolving logic of participant observation, educational ethnography, and other case studies. Review of Research in Education, 6, 316-377.
- Smith, M., & Glass, G. (1977). Meta-analysis of psychotherapy outcome studies. American Psychologist, 32, 752-760.
- Smith, R., Jensen, K., & Dillingofski, M. (1971). The effects of integrating reading and writing on four variables. Research in the Teaching of English, 5, 179-189.
- Smoke, T. (1991). Becoming an academic insider: One student's experience of attaining academic success in college (Doctoral dissertation, New York University, 1991). Dissertation Abstracts International, 52, 2055A.
- Song, B. (1992). A comparison of perceptions of EFL textbooks used in Chinese secondary schools (Doctoral dissertation, the University of Mississippi, 1991). Dissertation Abstracts International, 53, 90A.
- Söter, A. O. (1987). Recent research on writing: Implications for writing across the curriculum. Curriculum Studies, 19, 425-438.
- Spencer, E. (1983). Writing matters across the curriculum. London: Hodder and Stoughton.
- Spradley, J. P. (1980). Participant observation. New York: Holt, Rinehart and Winston.
- Spradley, J. P., & McCurdy, D. W. (1971). Conformity and conflict: Readings in cultural anthropology. Boston, MA: Little and Brown.
- Squire, J., & Applebee, R. (1968). High school English instruction today: The national study of high school English programs. New York: Appleton-Century-Crofts.
- Stallard, C. K., Jr. (1974). An analysis of the writing behavior of good student writers. Research in the Teaching of English, 8, 206-218.

- Stalnaker, J. M. (1933). Essay and objective writing tests. English Journal, 22, 222.
- Stalnaker, J. M. (1934). The construction and results of a twelve-hour test in English composition. School and Society, 39, 218-224.
- Starring, R. W. (1952). A study of ratings of comprehensive examination themes when certain elements are weakened (Doctoral dissertation, Michigan State College, 1952). Doctoral Dissertations accepted by American Universities, 12, 850.
- Stewart, M. F. (1978). Syntactic maturity from high school to university: A first look. Research in the Teaching of English, 12, 37-46.
- Stotsky, S. (1983a). Research on reading/writing relationships: A synthesis and suggested directions. Language Arts, 60, 627-642.
- Stotsky, S. (1983b). Types of lexical cohesion in expository writing: Implications for developing the vocabulary of academic discourse. College Composition and Communication, 34, 430-446.
- Stotsky, S. (1984). Research on reading/writing: A synthesis and suggested directions. In J. M. Jensen (Ed.), Composing and comprehending (pp. 7-31). Urbana, IL: National Council of Teachers of English.
- Stouffer, S. A. (1962). Social research to test ideas. New York: The Free Press.
- Stout, J. E. (1921). The development of high school curricula in the north central states from 1860 to 1918. Chicago: University of Chicago.
- Strickland, R. G. (1962, July). The language of elementary school children: Its relationship to the language of reading textbooks and the quality of reading of selected children. Bulletin of the School of Education, Indiana University, 38, pp. 1-131.
- Swanson-Owens, D. (1986). Identifying natural sources of resistance: A case study of implementing writing across the curriculum. Research in the Teaching of English, 20, 69-97.
- Tannen, D. (1982). Oral and literary strategies in spoken and written narratives. Language, 58, 1-21.

- Tate, G., & Corbett, E. P. J. (Eds.). (1970). Teaching high school composition. New York: Oxford University Press.
- Taylor, B., & Berkowski, S. (1980). Facilitating children's comprehension of content material. In M. Kamil, & A. Moe (Eds.), Perspectives on reading research and instruction (pp. 64-68), Twenty-Ninth Yearbook of the National Reading Conference. Washington DC: National Reading Conference, Inc.
- Taylor, H. (1981). Listening comprehension and reading comprehension as predictors of achievement on college composition (Doctoral dissertation, University of Washington, 1981). Dissertation Abstracts International, 42, 66A.
- Taylor, K. (1978). If not grammar, what?--Taking remedial writing instruction seriously. (ERIC Document Reproduction Service No. ED 159 668)
- Teng, C. (1991). Grammar, imitation, and process: How teaching methods affect what Chinese ESL students learn about written English (Doctoral dissertation, University of Kansas, 1990). Dissertation Abstracts International, 52, 452A.
- Thaiss, C. H., & Suhor, C. (Eds.). (1984). Speaking and writing K-12: Classroom strategies and the new research. Urbana, IL: National Council of Teachers of English.
- Thomas, F. (1976). The extent of the relationship between reading achievement and writing achievement among college freshmen (Doctoral dissertation, University of South Carolina, 1976). Dissertation Abstracts International, 37, 6320A.
- Thorndike, E. L. (1927). An inventory of English constructions with measures of their importance. Teachers College Record, 28, 580-610.
- Tierney, R. J., & Mosenthal, J. H. (1983). Cohesion and textual coherence. Research in the Teaching of English, 17, 215-229.
- Tippet, L. H. C. (1931). The methods of statistics. London: Williams and Norgate.
- Travers, M. W. (Ed.). (1973). Second handbook of research on teaching. Chicago: Rand McNally and Company.

- Toulman, S. E. (1958). The uses of argument. Cambridge, Britain: Cambridge University Press.
- Van Bruggen, J. A. (1946). Factors affecting regularity of the flow of words during written composition. Journal of Experimental Education, 40, 133-155.
- Vygotsky, L. S. (1962). Thought and language (F. Hanfmann & G. Vakar, Trans.). Cambridge: The M.I.T. Press. (Original work published 1934)
- Walker-Lewis, H. (1981). Using writing to improve the reading comprehension abilities of academically underprepared college students (Doctoral dissertation, Hofstra University, 1981). Dissertation Abstracts International, 43, 2053A.
- Wallace, W. (1971). The logic of science in sociology. Chicago: Aldine-Atherton, Inc.
- Walshe, R. (1981). Donald Graves in Australia. Sydney, Australia: PETA.
- Wang, C. (1992). Paragraph organization in English and Chinese academic prose: A comparative study (Doctoral dissertation, Indiana University of Pennsylvania, 1992). Dissertation Abstracts International, 53, 796A.
- Whyte, L. L., Wilson, A. G., & Wilson, D. (Eds.). (1969). Hierarchical structures. New York: Elsevier.
- Willms, J. D., & Raudenbush, S. W. (1989). A longitudinal hierarchical model for estimating school effects and their stability. Journal of Educational Measurement, 26, 209-232.
- Willson, V. L., & Putnam, R. R. (1982). A meta-analysis of pretest sensitization effects in experimental design. American Educational Research Journal, 19, 249-258.
- Wiseman, S. (1949). The marking of English composition in grammar school selection. British Journal of Educational Psychology, 19, 200-209.
- Wiseman, S., & Wrigley, J. (1958). Essay-reliability: The effect of choice of essay-title. Educational and Psychological Measurement, 18, 129-138.

- Witte, S. P., & Davis, A. (1982). The stability of T-unit length in written discourse of college freshmen: A second study. Research in the Teaching of English, 16, 71-93.
- Witte, S. P., & Faigley, L. (1981). Coherence, cohesion, and writing ability. College Composition and Communication, 32, 189-204.
- Wolf, F. M. (1986). Meta-analysis: Quantitative methods for research synthesis. Beverly Hills, CA: Sage.
- Wolfe, R. (1975). An examination of the effects of teaching a reading vocabulary upon writing vocabulary in student composition (Doctoral dissertation, University of Maryland, 1975). Dissertation Abstracts International, 36, 3337A.
- Woodfin, M. (1968). Correlations among certain factors and the written expression of third grade children. Educational and Psychological Measurement, 28, 1237-1242.
- Woodward, J. C., & Phillips, A. (1967). Profile of the poor writer--The relationship of selected characteristics to poor writing in college. Research in the Teaching of English, 1, 41-53.
- Xu, M. (1990). Assessment of international graduate students' perceived language needs, proficiency, and academic performance (Doctoral dissertation, State University of New York at Albany, 1990). Dissertation Abstracts International, 51, 1538A.
- Young, R. E., & Becker, A. L. (1965). Toward a modern theory of rhetoric: A tagmemic contribution. Harvard Educational Review, 35, 64-66.
- Zeman, S. (1969). Reading comprehension and writing of second and third graders. Reading Teacher, 23, 144-150.

APPENDICES

APPENDIX A

SELECTED BIBLIOGRAPHY AS CITED IN THE TWENTY REPORTS: CLASSIFIED

Audience Effect on Writing

Berkenkotter, C. (1981). Understanding a writer's awareness of audience. College Composition and Communication, 32, 388-399. (RM2)

Bracewell, R. J., Scardamalia, M., & Bereiter, C. (1978). The development of audience awareness in writing. Paper presented at the annual meeting of American Educational Research Association, Toronto, Canada. (RM2)

Crowhurst, M., & Piche, G. L. (1979). Audience and mode of discourse effects on syntactic complexity in writing at two grade levels. Research in the Teaching of English, 18, 383-417. (RM3, RN)

Nold, E. W., & Freedman, S. W. (1977). An analysis of readers' responses to essays. Research in the Teaching of English, 11(12), 164-174. (RK)

Rubin, D. L., & Piche, G. L. (1979). Development in syntactic and strategic aspects of audience adaptation skills in written persuasive communication. Research in the Teaching of English, 13, 293-316. (RM2)

Classroom Strategies

Bereiter, C., & Scardamalia, M. (1978). From conversation to composition. In R. Glaser (Ed.), Advances in instructional psychology (pp. 10-12). Hillsdale, NJ: Erlbaum. (RO)

Bridge, C., Hiebert, E. H., & Chesky, J. (1983). Classroom writing practices. In J. A. Niles and L. A. Harris (Eds.), Searches for meaning in reading/language processing and instruction (pp. 238-242). Rochester, NY: National Reading Conference, Inc. (RM2)

Britton, J. (Ed.) (1967). Talking and writing: A handbook for English teachers. London: Methuen. (RBE)

Brostoff, A. (1979). Good assignments lead to good writing. Social Education, 43, 184-86. (RHA)

- Buckley, M. H., & Boyle, O. (1983). Mapping and composing. In M. Myers and J. Grey (Eds.), Theory and practice in the teaching of composition (pp. 59-66). Urbana, IL: National Council of Teachers of English.
- Cooper, C. R. (1973). Students learn to read fiction by writing it. California English Journal, 9, 7-17. (RD1)
- Daiker, D., Kerek, A., & Morenberg, M. (Eds.). (1979). Sentence-combining and the teaching of writing. Conway: University of Central Arkansas. (RI)
- Kitzhaber, A. R. (1963). Themes, theories, and therapy: The teaching of writing in college. New York: McGraw-Hill. (RA)
- Moffet, J. W. (1968b). A student-centered language arts curriculum, grades K-13: A handbook for teachers. New York: Houghton Mifflin. (RD1)
- Murray, D. M. (1968). A writer teaches writing: A practical method of teaching composition. Boston, MA: Houghton Mifflin. (RBE)
- Ney, J. W. (1966). Applied linguistics in the seventh grade. English Journal, 55, 895-897. (RBO)
- Small, R., Jr. (1979). The YA novel in the composition program, Part II. English Journal, 68(7), 75-77. (RHA)
- Tate, G., & Corbett, E. P. J. (Eds.). (1970). Teaching high school composition. New York: Oxford University Press. (RBE)
- Thaiss, C. H., & Suhor, C. (Eds.). (1984). Speaking and writing K-12: Classroom strategies and the new research. Urbana, IL: National Council of Teachers of English. (RM3)

Cognitive Theory

- Crowhurst, M. (1983a). Persuasive writing at grads 5, 7, and 11: A cognitive-developmental perspective. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada. (RO)
- Flower, L. (1979). Writer-based prose: A cognitive basis for problems in writing. College English, 41, 19-37. (RM3)

- Flower, L., & Hayes, J. (1980a). The dynamics of composing: Making plans and juggling constraints. In L. W. Gregg and E. R. Steinberg (Eds.), Cognitive processes of writing (pp. 31-50). Hillsdale, NJ: Lawrence Erlbaum. (RM2, RM3)
- Flower, L., & Hayes, J. (1980b). Cognition of discovery: Defining a rhetorical problem. College Composition and Communication, 31, 21-32. (RK, RM2)
- Flower, L., & Hayes, J. (1977). Problem solving strategies and the writing process. College English, 39, 449-461. (RM3)
- Gregg, L. W., & Steinberg, E. R. (Eds.). (1980). Cognitive processes of writing. Hillsdale, NJ: Lawrence Erlbaum. (RM3)
- Healy, M. A. K. (1981). Purpose in learning to write: An approach to writing in three curricular areas. In C. H. Fredericksen and J. F. Dominie (Eds.), Writing: the nature, development, and teaching of written communication: Vol. 2. Processes, development and communication (pp. 223-233). Hillsdale, NJ: Lawrence Erlbaum. (RM3)
- Purves, A. C., & Tierney, R. (1983). On the nature of rhetorical communities. Urbana: Curriculum Laboratory, University of Illinois at Urbana-Champaign. (RM3)
- Scardamalia, M., & Bereiter, C. (1986). Written composition. In M. Wittrock (Ed.), Handbook of research on teaching (3rd ed.) (pp. 778-803). New York: Macmillan. (RM3)
- Vygotsky, L. S. (1962). Thought and language (F. Hanfmann & G. Vakar, Trans.). Cambridge: The M.I.T. Press. (Original work published 1934) (RL)

Computer Assisted Writing

- Iker, H. P., & Harway, N. (1965) A computer approach towards the analysis of content. Behavioral Science, 10, 173-182. (RBE)
- Page, E. B. (1968). The use of the computer in analyzing student essays. International Review of Education, 14, 253-263. (RF)
- Page, E. B., & Paulus, D. H. (1968). The analysis of essays by computers. Storrs: University of Connecticut. (RBE)

Content Knowledge on Writing

Hilgers, T. L. (1982). Experimental control and the writing stimulus: The problem of unequal familiarity with content. Research in the Teaching of English, 16, 381-190. (RM2)

Langer, J. A. (1984). The effects of available information on responses to school writing tasks. Research in the Teaching of English, 18, 27-44. (RM2)

Discourse Theory

Britton, J. (1982). Shaping at the point of utterance. In G. Pradl (Ed.), Prospect and retrospect: Selected essays of James Britton (pp. 139-145). London: Heinemann. (RM3)

Connors, R. A. (1981). The rise and fall of the modes of discourse. College Composition and Communication, 32, 444-455. (RK)

Crowhurst, M. (1981). Cohesion in argumentative prose written by sixth-, tenth-, and twelfth-graders. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles. (RO)

De Beaugrande, R. (1980). Text, discourse and process: Toward a multidisciplinary science of texts. Norwood, NJ: Ablex Publishing Corporation. (RK)

De Beaugrande, R., & Dressler, W. (1981). Introduction to text linguistics. New York: Longman. (RK)

Halliday, M. A. K., McIntosh, A., & Stevens, P. (1964). The linguistic sciences and language teaching. Bloomington: Indiana University Press. (RBE)

Jacobson, L. (1969). Rhetoric and stylistics: Some basic issues in the analysis of discourse. College Composition and Communication, 20, 314-328. (RBE)

Kineavy, J. E. (1971). A theory of discourse. Englewood Cliffs, NJ: Prentice-Hall. (RK)

Kineavy, J. E. (1969). The basic aims of discourse. College Composition and Communication, 20, 297-304. (RBE)

Moffet, J. (1968a). Teaching the universe of discourse. Boston, MA: Houghton Mifflin. (RBE, RM3, RO)

- Postman, N. (1967). Linguistics and the pursuit of relevance. English Journal, 56, 1160-1165. (RBO)
- Reed, W. M., Burton, J. K., & Kelly, P. P. (1985). The effects of writing ability and mode of discourse on cognitive capacity engagement. Research in the Teaching of English, 19, 283-297. (RN)
- Toulman, S. E. (1958). The uses of argument. Cambridge: Cambridge University Press. (RO)

Editing

- Bridwell, L. S. (1980). Revising strategies in twelfth grade students' transactional writing. Research in the Teaching of English, 14, 197-222. (RJ)
- Crowhurst, M. (1983b). Revision strategies of students at three grade levels. (Final report). Educational Research Institute of British Columbia. (RO)

Evaluation and Measurement of Writing

- Anastasi, Anne. (1968). Psychological testing (3rd ed.). New York: Macmillan. (RD1)
- Anderson, C. C. (1960). The new STEP Essay Test as a measure of composition ability. Educational and Psychological Measurement, 20, 95-102. (RA)
- Bamberg, B. (1983). What makes a text coherent? College Composition and Communication, 34, 417-429. (RK)
- Beach, R. (1976). Self-evaluation strategies of extensive revisers and nonrevisers. College Composition and Communication, 27, 160-164. (RF)
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). Handbook of formative and summative evaluation of student learning. New York: McGraw-Hill. (RD1)
- Cast, B. M. D. (1940). The efficiency of different methods of marking English composition. British Journal of Educational Psychology, 9, 257-269. (RA)
- Cooper, C., & Odell, L. (1977). Evaluating writing: Describing, meaning, judging. Urbana, IL: National Council of Teachers of English. (RM1, RM3)
- Diederich, P. (1974). Measuring growth in English. Urbana, IL: National Council of Teachers of English. (RF)

- Diederich, P. (1965). Grading and measuring. In A. Jewett (Ed.), Improving english composition (pp. 81-91). Washington, DC: NEA. (RD1)
- DiStefano, P., & Howie, S. (1979). sentence weights: An alternative to the T-unit. English Education, 11, 98-101. (RM1)
- Fahnestock, J. (1983). Semantic and lexical coherence. College Composition and Communication, 34, 400-416. (RK)
- Gillie, P. J. (1957). A simplified formula for measuring abstraction in writing. Journal of Applied Psychology, 41, 214-217. (RA)
- Halliday, M. A. K., & Hasan, R. (1976). Cohesion in English. London: Longman. (RK, RO)
- Hartog, P., & Rhodes, E. C. (1936). The marks of examiners. London: Macmillan Company. (RA)
- Huddleson, E. M. (1954). Measurement of writing ability at the college-entrance level: Objective vs. subjective testing techniques. Journal of Experimental Education, 22, 165-213. (RA)
- Hunt, K. (1976). A synopsis of clause-to-sentence length factors. In R. Graves (Ed.), Rhetoric and composition: A sourcebook for teachers (pp. 110-117). Rochelle Park, NJ: Hayden. (RN)
- Hunt, K. (1970). Syntactic maturity in school children and adults. Monographs of the Society for Research in Child Development, 35(Serial No. 134). (RE, RM1, RN)
- Markels, R. B. (1983). Cohesion paradigms in paragraphs. College English, 45, 450-464. (RK)
- McColly, W. (1970). What does educational research say about the judging of writing ability. Journal of Educational Research, 64, 148-156. (RD1)
- Mullis, I. V. S., & Mellon, J. C. (1980). Guideline for describing three aspects of writing: Syntax, cohesion, and mechanics. Denver, CO: National Assessment of Educational Progress. (ERIC Document reproduction Service No. ED 145 482) (RM2)
- Myers, M. (1980). A procedure for writing assessment and holistic scoring. Urbana, IL: National Council of Teachers of English. (RM2)

- Stalnaker, J. M. (1934). The construction and results of a twelve-hour test in English composition. School and Society, 39, 218-224. (RA, RBO)
- Stalnaker, J. M. (1933). Essay and objective writing tests. English Journal, 22, 222. (RA)
- Stotsky, S. (1983b). Types of lexical cohesion in expository writing: Implications for developing the vocabulary of academic discourse. College composition and Communication, 34, 430-446. (RK)
- Tierney, R. J., & Mosenthal, J. H. (1983). Cohesion and textual coherence. Research in the Teaching of English, 17, 215-229. (RK)
- Witte, S. P., & Davis, A. (1982). The stability of T-unit length in written discourse of college freshmen: A second study. Research in the Teaching of English, 16, 71-93. (RN)
- Witte, S. P., & Faigley, L. (1981). Coherence, cohesion, and writing ability. College Composition and Communication, 32, 189-204. (RK)

Prewriting

- Rohman, D. G. (1965). Pre-writing: The stage of discovery in the writing process. College Composition and Communication, 16, 106-112. (RBE)
- Scardamalia, M. (1982). Assimilative processes in composition planning. Educational Psychologist, 17, 165-171. (RM3)

Research Methodology

- Anderson, R. C. (1977). The notion of schemata and the educational enterprise. In P. C. Anderson, P. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge (pp. 415-431). Hillsdale, NJ: Erlbaum. (RM2)
- Campbell, D., & Stanley, J. (1963). Experimental and quasi-experimental designs for research. Chicago: Rand McNally College Publishing Co. (RF)
- Gioconia, R., & Hedges, L. V. (1982). Identifying features of effective open education. Review of Educational Research, 52, 579-602. (RJ)

- Glass, G. V. (1978). Integrating findings: the meta-analysis of research. In L. S. Shulman (Ed.), Review of research in education (5th ed.), (pp. 351-379). Itasca, IL: F. E. Peacock. (RJ)
- Hedges, L. V. (1982a). Estimating effect size from a series of independent experiments. Psychological Bulletin, 92, 490-499. (RJ)
- Hedges, L. V. (1982b). Fitting categorical models to effect sizes from a series of experiments. Journal of Educational Statistics, 7, 119-137. (RJ)
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. Journal of Educational Statistics, 6, 107-128. (RJ)
- Huber, P. J. (1977). Robust statistical procedures. Philadelphia: Society for Industrial and Applied Mathematics. (RJ)
- Kerlinger, F. N. (1973). Foundations of behavioral research. New York: Holt, Reinhart, and Winston. (RN)
- Nunnally, J. C. (1967). Psychometric theory. New York: MacGraw-Hill. (RN)

Rhetoric

- Bailey, D. (Ed.). (1965). Essays on rhetoric. New York: Oxford University Press. (RBE)
- Bain, A. (1867). English composition and rhetoric (2nd English ed.). New York: D. Appleton and Company. (RK)
- Connolly, F. X. (1953). A rhetoric casebook. New York: Harcourt Brace Jovanovich. (RBE)
- Corbett, E. P. J. (1965). Classical rhetoric for the modern student. New York: Oxford University Press. (RBE)
- Martin, H., Ohmann, R., & Wheatley, J. (1969). The logic and rhetoric of exposition. New York: Holt, Rinehart and Winston. (RBE)
- Rockas, L. (1964). Modes of rhetoric. New York: St. Martin's Press. (RBE)
- Young, R. E., & Becker, A. L. (1965). Toward a modern theory of rhetoric: A tagmemic contribution. Harvard Educational Review, 35, 64-66. (RBO)

Writer

Smith, F. (1981). Writing and the writer. New York: Holt, Rinehart & Winston. (RM3)

Walshe, R. (1981). Donald Graves in Australia. Sydney, Australia: PETA. (RM3)

Writing Across the Curriculum

Fulwiler, T. E. (1984). How well does writing across the curriculum work? College English, 46, 113-126. (RL)

Martin, J. (1976). Writing and learning across the curriculum. London: Ward Lock Education. (RL)

Newell, G. (1984). Learning from writing in two content areas: A case study/protocol analysis. Research in the Teaching of English, 18, 265-287. (RM2)

Spencer, E. (1983). Writing matters across the curriculum. London: Hodder and Stoughton. (RM3)

Writing Apprehension

Daly, J., & Hexamer, A. (1983). Statistical power in research in English education. Research in the Teaching of English, 17, 156-164. (RN)

Daly, J., & Miller, M. D. (1975). The empirical development of an instrument to measure writing apprehension. Research in the Teaching of English, 9, 272-289. (RF, RN)

Daly, J., & Wilson, D. A. (1983). Writing apprehension, self-esteem, and personality. Research in the Teaching of English, 17, 327-342. (RN)

Faigley, L., Daly, J., & Witte, S. (1981). The effects of writing apprehension on writing performance and competence. Journal of Educational Research, 75, 16-21. (RM2, RN)

Reed, W. M., Vandett, N. M., & Burton, J. K. (1984). The effects of writing skills, sex, and writing preparation on writing apprehension. FORUM in Reading and Language Education, 1, 45-60. (RN)

Writing Process

Beach, R., & Bridwell, L. (Eds.). (1984). New directions in composition research. New York: Guilford Press. (RM3)

Emig, J. (1971). The composing processes of twelfth graders. Urbana, IL: National Council of Teachers of English. (RC, RD1, RF, RM3)

Perl, S. (1979). The composing processes of unskilled college writers. Research in the Teaching of English, 13, 317-336. (RM3)

Writing Programs

Camp, G. (Ed.). (1982). Teaching writing: Essays from the Bay Area Writing Project. Berkeley, CA: Boynton-Cook. (RM3)

Writing Research

Bateman, D. R., & Zidonis, F. J. (1964). The effect of a knowledge of generative grammar upon the growth of language complexity. Columbus: The Ohio State University. (RC)

Bereiter, C. (1960). Verbal and ideational fluency in able 10th graders. Journal of Educational Psychology, 51, 337-345. (RBE)

Bereiter, C., & Scardamalia, M. (1983). Levels of inquiry in writing research. In P. Mosenthal et al. (Eds.), Research on writing: Principles and methods (pp. 20-23). London: Longman. (RM3)

Braddock, R., Lloyd-Jones, R., & Schoer, L. (1963). Research in written composition. Champaign, IL: National Council of Teachers of English. (RBE, RC, RD1, RJ)

Britton, J., Burgess, T., Martin, N., McLeod, A., & Rosen, H. (1975). The development of writing abilities (11-18). London: Macmillan Education. (RH2, RM2)

Burton, D. L. (1973). Research in the teaching of English: the troubled dream. Research in the Teaching of English, 7, 160-189. (RJ)

Christensen, F. (1967). Notes toward a new rhetoric: Six essays for teachers. New York: Harper & Row. (RBE, RC, RD1, RJ, RM1)

Christiansen, M. A. (1965). Tripling writing and omitting readings in freshman English: An experiment. College Composition and Communication, 16, 122-124. (RI)

- Cooper, C., & Odell, L. (Eds.). (1978). Research on composing: Points of departure. Urbana, IL: National Council of Teachers of English. (RF)
- Dillon, G. L. (1981). Constructing texts: Elements of a theory of composition and style. Bloomington: Indiana University Press. (RM2)
- Donelson, K. L. (1967). Variables distinguishing between effective and ineffective writers at the tenth grade. Journal of Experimental Education, 35, 37-41. (RI)
- Emig, J. (1983). The web of meaning: Essays on writing, teaching, learning, and thinking. Upper Montclair, NJ: Boynton/Cook. (RL)
- Emig, J. (1982). Inquiry paradigms and writing. College Composition and Communication, 33, 64-75. (RJ)
- Emig, J. (1977). Writing as a mode of learning. College Composition and Communication, 28, 122-127. (RL)
- Graves, D. H. (1980). Research update: A new look at writing research. Language Arts, 57, 913-919. (RJ)
- Hays, J. N., Roth, P. A., Ramsey, J. R., & Foulke, R. D. (Eds.). (1983). The writer's mind: Writing as a mode of thinking. Urbana, IL: National Council of Teachers of English. (RL)
- Hillocks, G., Jr. (1981). The response of college freshmen to three modes of instruction. American Journal of Education, 89, 373-395. (RJ)
- Hunt, K. (1965). Grammatical structures written at three grade levels. Urbana, IL: National Council of Teachers of English. (RBE, RD1, RE, RK, RM1)
- Langer, J., & Applebee, A. (1986). Writing and learning in the secondary school, Final Report to the National Institute of Education. Stanford, CA: Stanford University. (RL)
- Langer, J., & Applebee, A. (1982). Writing and learning in the secondary school, Proposal to the National Institute of Education. Stanford, CA: Stanford University. (RL)
- Meckel, H. C. (1963). Research on teaching composition and literature. In N. L. Gage (Ed.), Handbook of research on teaching (pp. 966-1006). Chicago: Rand McNally and Co. (RBE, RI)

Mellon, J. C. (1969). Transformational sentence-combining: A method for enhancing the development of syntactic fluency in English composition (Research Report No. 10). Urbana, IL: National Council of Teachers of English. (RBE, RBO, RC, RJ, RM1)

Meyers, M. (Ed). (1985). The teacher-researcher: How to study writing in the classroom. Urbana, IL: National Council of Teachers of English. (RM3)

O'Hare, F. (1971). Sentence-combining: Improving student writing without formal grammar instruction (Research Report No. 15). Urbana, IL: National Council of Teachers of English. (RC, RD1, RJ)

Parker, R. P. (1979). From Sputnik to Dartmouth: Trends in the teaching of composition. English Journal, 68, 32-37. (RJ)

Travers, M. W. (Ed.). (1973). Second handbook of research on teaching. Chicago: Rand McNally and Co. (RI)

Van Bruggen, J. A. (1946, December). Factors affecting regularity of the flow of words during written composition. Journal of Experimental Education, 40, 133-155. (RA, RBE)

Wiseman, S., & Wrigley, J. (1958). Essay-reliability: The effect of choice of essay-title. Educational and Psychological Measurement, 18, 129-138. (RA)

Writing Survey

Applebee, A. N. (1984). Contexts for learning to write: Studies of secondary school instruction. Norwood, NJ: Ablex. (RL, RM3)

Applebee, A. N. (1982). Writing and learning in school setting. In M. Nystrand (Ed.), What writers know (pp. 365-381). New York: Academic Press. (RM3)

Applebee, A. N. (1981). Writing in the secondary school. (Research Report No. 21). Urbana, IL: National Council of Teachers of English. (RJ, RL, RM2, RO)

Applebee, A. N., Langer, J. A., & Mullis, I. V. (1986). The writing report card: Writing achievement in American schools. Princeton, NJ: Educational Testing Services. (RO)

Squire, J., & Applebee, R. (1968). High school English instruction today: The national study of high school English programs. New York: Appleton-Century-Crofts. (RF)

Writing/Culture

Cole, M., & Scribner, S. (1974). Culture and thought: A psychological introduction. New York: Scribner's. (RM3)

Purves, A. C., & Purves, W. (1986). Cultures, text models and the activity of writing: Speculations for future research. Urbana, IL: Curricular Laboratory, University of Illinois at Urbana-Champaign. (RM3)

Writing/Reading

Evans, R. (1979). The relationship between the reading and writing of syntactic structures. Research in the Teaching of English, 13, 129-136. (RI)

Jensen, J. M. (Ed.). (1984). Composing and comprehending. Urbana, IL: National Council of Teachers English. (RM3)

Meyer, B. (1982). Reading research and the composition teacher: The importance of plans. College Composition and Communication, 33, 37-49. (RI)

Purves, A. C. (1983). Language and processing: Reading and writing. College English, 45, 129-140. (RM2)

Smith, R., Jensen, K., & Dillingofski, M. (1971). The effects of integrating reading and writing on four variables. Research in the Teaching of English, 5, 179-189. (RI)

APPENDIX B

Survey: Using Writing Assignments in the Classroom

PERSONAL INFORMATION:

1. The grade level you teach. (If you teach more than one grade level, select one level to be the basis for this survey.)

2. The subject you teach. (If you teach more than one subject, select one subject to be the basis for this survey.)

3. Year of teaching experience. (Exclude student teaching.)

4. Male _____ Female _____

=====

Please respond to the following questions by placing check marks in the appropriate boxes. Please limit yourself to (1) one grade level, (2) one subject, and (3) one intact classroom of students. Also, exclude essay examinations and answers to questions when defining writing assignments.

1. How long are the writing assignments you give your students? (Check as many squares as appropriate.)
 - a. under 100 words []
 - b. between 100 and 300 words []
 - c. between 300 and 500 words []
 - d. between 500 and 1,500 words []
 - e. over 1,500 words []
2. Over a period of a school year, how many writing assignments do you give? (check only one square.)
 - a. 1-3 []
 - b. 4-6 []
 - c. 7-9 []
 - d. 10-12 []
 - e. 13 or more []
3. What types of writing do you assign? (Check as many squares as appropriate.)
 - a. narration (telling stories, anecdotes, personal experiences) []
 - b. exposition (explaining) []
 - c. argumentation []

- d. reporting []
4. What is the basis for assigning writing? (Check as many squares as appropriate.)
- a. a summary of what has been covered in class []
 - b. an extension and/or expansion of what has been covered in class []
 - c. a substitute for what cannot be covered in class []
5. How do you teach writing, with respect to your assignment? (Check as many squares as appropriate.)
- a. by an explanatory assignment sheet []
 - b. by using a model paper for student examination []
 - c. having students write assignments in class, or at least partly in class, under your supervision []
 - d. breaking the assignment into steps or stages and teaching each step separately []
 - e. using student editing committees to offer helpful suggestions for fellow students in the process of writing []
6. What types of corrections do you make on student papers? (Check as many squares as appropriate.)
- a. indicating errors in spelling, punctuation, grammar, usage, and manuscript appearance, but not correcting the errors []
 - b. indicating errors in spelling, punctuation, grammar, usage, and manuscript appearance, and correcting the errors []
 - c. indicating faulty sentences, e.g., vagueness, ambiguity, lack of sense--but not rewriting the sentences []
 - d. indicating faulty sentences and rewriting them []
7. What types of comments do you make on the papers you assign? (Check as many squares as appropriate.) Comments not to be confused with corrections.
- a. comments about form (manuscript appearance, grammar, spelling, punctuation) written in the margins []
 - b. comments about form written in the summary statement []
 - c. comments about content written in the margins []
 - d. comments about content written in the summary statement []

8. What is the basis for your evaluation of the assignments? (Check as many squares as appropriate.)
- a. evaluation based on content only []
 - b. evaluation based on form only []
 - c. evaluation based on a combination of form and content (If you check this item, answer the following three subquestions.) []
 - 1. equal emphasis on form and content []
 - 2. more emphasis on form than on content []
 - 3. more emphasis on content than on form []
9. How are grades assigned on the papers? (Check as many squares as appropriate.)
- a. a grade appears on the paper with no evaluative comments []
 - b. a grade appears on the paper together with evaluative comments []
 - c. evaluative comments appear on the paper with no grade assigned []
10. After students receive their papers back, they
- a. keep them in their own personal files, located in their notebooks []
 - b. keep them in classroom files []
 - c. handle them in no systematic way []
 - d. resubmit them in revised form for further evaluation []
11. The teaching of writing should be the responsibility of
- a. the English teacher []
 - b. the content area teacher []
 - c. college or university personnel []

(Source: Donlan, 1974, 252-254)

APPENDIX C

Examples of Exercises Used in NAEP

I. Acceptable/Unacceptable Exercises:

Letter of Invitation (Age 13: no. 3)

About three months ago, Leo Logan moved from the city to the county. His father bought a farm, and now Leo's address is Rebel Road, Rural Delivery No. 1, Harris, Nebraska 69000.

Leo likes the country, but he misses his old friend Ozzie Drake. Leo's mother says, "Why don't you write to Ozzie and invite him to visit us for a week this summer?"

Write Leo's letter of invitation to Ozzie.

[Approximately 1-1 1/2 pages of lined space were provided for the response.]

Recording a Telephone Message (Age 13: no. 4)

You are going to hear a telephone conversation between two boys, Al and Ben. During the conversation, you will discover that Ben is going to have to write a note to his mother. Listen carefully to find out the things that Ben will have to say in his note. [children then listened to the conversation reproduced below. It was not printed in the children's booklet.]

Al: Hey, Ben, this is Al. I called you to remind you you're supposed to come down to my house for supper tonight before the game.

Ben: Oh yeah, I remember. Your mother and father are going to pick me up on the way back from the barber shop. What time do you think you'll get here?

Al: Oh, around 5 o'clock.

Ben: O.K. I'd better leave a note for my mother. She won't be home until 5:30 and maybe she's forgotten that I won't be eating at home.

Al: You'd better remind her of the ball game,

too. She's supposed to pick us up afterwards, isn't she?

Ben: That's right. Your father is taking us to the game but my mother is bringing us home. Thanks for reminding me. I'll write the note right away.

You will now have another chance to hear the conversation between Al and Ben. Listen to it carefully again and then, when the boys have finished talking, write the note that you think Ben should leave for his mother.

[Approximately 1/2 page of lined space was provided for the response.]

II. Yes/No Exercises:

1. Are there jobs requiring much writing?
2. Have you ever sent writing in for publication?
3. During the past year have you
left someone a written message?
written a letter ordering a product by mail?
4. Other than in school have you ever written
a song lyric?
a story?
a poem?
a joke?
5. Other than in school have you ever written
a play?
a report?
a newspaper story?
a magazine article?
6. If on a trip last year, did you write
a thank-you note to someone visited?
a letter about the trip?
notes on the places seen?
a post card?
a letter while on trip?
a report about the trip?
a trip diary?

Essay Exercises:

Essay Topics for Age 9
Going to School

Think about what happens when you go to school. Write a little story that tells what you do from the time you leave where you live until you get to school. Be sure to include everything that you think is important.

Forest Fire

Here is a picture of something sad that is going on in the forest. Look at the picture for a while. Do you see the forest fire? Write a story about what is happening in the picture. This is an important story because you want people to know about this sad event.

[The picture depicts a forest fire with animals swimming across a river rapids to obtain safety.]

Astronaut

Here is a picture of an astronaut on the moon. Look at the picture for a while and think about what is happening. Now, pretend that YOU were the astronaut, and write a story about your walk on the moon.

Essay Topics for Age 13

Pen-pal Letter

Write a friendly letter to a "pen-pal" in another country, telling him (or her) about your preparations for some holiday. Your letter should be as "newsy" as possible because your friend probably does not know how you do things where you live.

Historical Event

Most of us look up to some famous persons as a representative of the things we believe in or as the kind of person we would like to be. This person may come from any part of our society. For instance, we might admire Winston Churchill or Martin Luther King, Walter Schirra or Mickey Mantle, Florence Nightingale or Barbra Streisand. No matter where this person comes from or what kind of work he or she does, however, we can recognize such traits or greatness as determination, physical courage, the ability to inspire others, and faithfulness to some worthy cause. Think about a famous person whom you admire. Select a particularly admirable characteristic or quality of that person--such as Mickey

Mantle's courage in the face of crippling physical handicaps or Florence Nightingale's determination to fight against strong governmental pressure. Write an essay of about 200-250 words describing this characteristic or quality. Be sure to provide an illustration of it from the person's life. Try to show that the person is great at least partly because of this characteristic or quality.

Essay Topics for Age 17

Tomato Lady

Here is a picture of a woman with some tomatoes. Look at the picture for a while and decide what is going on. When you have decided, write a story that tells what is happening in the picture and what is likely to happen next.

Famous Person

[Topic identical to that used at age 13.]
(16-22)

APPENDIX D

Mellon's study (1975) listed the revised objectives for writing as follows:

Writing

1. Demonstrate ability in writing to reveal personal feelings and ideas:
 - a. through free expression;
 - b. through the use of conventional modes of discourse.
2. Demonstrate ability to write in response to a wide range of societal demands and obligations (ability is defined to include correctness in usage, punctuation, spelling, and form or convention as appropriate to particular writing tasks, e.g., manuscripts, letters):
 - a. social
 1. personal
 2. organizational
 3. community
 - b. business/vocational
 - c. scholastic
3. Indicate the importance attached to writing skills:
 - a. recognize the necessity of writing for a variety of needs (as in 1 and 2);
 - b. Write to fulfill those needs;
 - c. get satisfaction, even enjoyment, from having written something well. (102)

VITA

Ruixuan Mao was born in Shanghai, China, on September 22, 1944. He attended elementary school there and, later on, moved to Chengdu, Sichuan, where he finished his secondary education. He graduated from high school in 1960 and entered the Southwest China Teachers University in the same year, majoring in English. In 1964, he graduated from the University and became a teacher of English there. In 1980, he was promoted to the position of lecturer and in 1986 to that of associate professor.

Ruixuan Mao entered the Graduate School of the University of Tennessee, Knoxville in 1986. He obtained his M. S. degree in Curriculum and Instruction with a major in English Education in 1986. In December, 1993, he was awarded the degree of Doctor of Philosophy in Education with a major in Curriculum and Instruction and a minor in English Education at the University of Tennessee, Knoxville. In 1988, while he was writing his dissertation, he was awarded the Hilton A. Smith Fellowship by the Graduate School of the University of Tennessee, Knoxville. In 1989, he was appointed as instructor of Chinese and put in charge of the Chinese language program at the University of Tennessee, Knoxville. He is a member of AERA and of NCTE and a member of several of China's academic associations of English, literature, translation and foreign studies at various levels as well.

Ruixuan Mao is married to Mi Hu. They have two sons. In August, 1992, Ruixuan Mao became Assistant Professor of English at Philander Smith College in Little Rock, Arkansas, where he teaches Freshman Composition, Advanced Composition, History of the English Language, Contemporary Grammar and Methodology for Secondary English. He also serves on the Teacher Education Committee of Philander Smith College.

