THE RELATIONSHIP BETWEEN CRITICAL THINKING AND EXPOSURE OF HIGH SCHOOL STUDENTS TO SCIENCE COURSES

A Thesis

Presented to

the Faculty of Graduate Studies

The University of Manitoba

In Partial Fulfillment

of the Requirements for the Degree

Master of Education

Ralph H. Berry

May, 1978

THE RELATIONSHIP BETWEEN CRITICAL THINKING AND EXPOSURE OF HIGH SCHOOL STUDENTS TO SCIENCE COURSES

BY

RALPH H. BERRY

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF EDUCATION

ABSTRACT

The problem of this study was to examine the growth of critical thinking of senior high school students as a function of the number and kind of science courses that they completed in grades eleven and twelve.

A sample of two hundred twenty students was tested with the Watson-Glaser Critical Thinking Appraisal Form YM at the end of grade ten. By the end of grade twelve, the sample size had been reduced, for reasons beyond the control of the experimenter, to sixty two students. These students were tested using the Watson-Glaser Critical Thinking Appraisal Form ZM.

After data collection, the gain in critical thinking was expressed in terms of gain between pretest and posttest z-scores. Then, correlational and anlysis of variance techniques were applied. In this analysis, the subjects were classified on the bases of age, sex, class (science major, science minor and non-science) and type (university entrance, mixed and general). The effects of specific subjects such as physics, biology, chemistry, and mathematics were examined. Within the limitations of the small sample in the study, the importance of science courses as a factor in the growth of critical thinking in grades eleven and twelve was established.

The results of the study point to several conclusions which merit discussion and from which implications for future educational practice may be drawn.

ACKNOWLEDGMENTS

The writer wishes to thank Dr. H. D. Williams for his patient and constructive guidance in the completion of this thesis, Dr. H. Grunau for his helpful suggestions and encouragement and also Dr. G. Hoare for many constructive criticisms.

Among others to whom he owes a debt of gratitude are Superintendent L. Milne and Principal G. Sefton for permission to carry out the study in the schools; to Messrs. John Fotheringhame and Terry Robertson for computer processing of data and to Dr. W. Clifford Carbno for his suggestions about the statistical analysis.

R. H. B.

Winnipeg, Manitoba.
May, 1978.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table		Page

CHAPTER I

INTRODUCTION

Importance of the Problem

Some thirty-five years ago, Dressel (1955) underscored the importance of the development of critical thinking as an educational goal:

> Critical thinking then is evidently the desired integrating principle or goal of education, the achievement of which promises that there will be a life-long interest in learning.

This stated importance of critical thinking has been emphasized in the development of science courses. Such courses are designed to foster in students the ability to think critically and to help students evaluate data gained scientifically while remaining open-minded and tolerant of the opinion of others. For example,

> A course in Physics should serve a more valuable purpose than simply the acquiring of basic information in the form of facts, principles and formulae. It should primarily:
>
> 1. Provide an understanding of scientific principles...
>
> 2. Be designed to develop the student's ability to visualize relationships...
>
> 3. Provide the student with an opportunity to do some individual experimentation, to develop his powers of observation and increase his ability to gather, interpret and analyze data independently. Such experience should enable him to acquire some facility in handling scientific apparatus, providing opportunity to test the accuracy of his measurements, to draw generalizations from his results, and become acquainted to some small degree with the experimental technique of scientific investigation, (Manitoba High School Program of Studies, 1967, p.42).

The question arose then of the effectiveness of the new science courses, as currently taught, in reaching this objective. Hence, there was a need for a study to evaluate the effectiveness of science courses per se and of science courses as opposed to non-science courses. The approach selected was to compare the performance in terms of gain scores of each of these classes of students.

The Problem of the Study

The problem of this study was to examine the growth in critical thinking of senior high school students as a function of the number and kind of science courses that they completed in grades eleven and twelve.

The Delineation of the Study

Since 1960, a number of science courses have been adopted in the secondary education programs in Manitoba Schools. These courses, including I.P.S., CHEM Study, PSSC Physics and BSCS Biology, were designed to emphasize the structure and process of science through scientific enquiry. Indeed, the Program of Studies for Science Grades 7-9, as authorized by the Minister of Education for the Province of Manitoba in 1967, emphasized this objective as follows:

> A major goal of the discovery approach is the development of intellectual independence. It seeks to raise thinking above mere memorization and recall. Pupils are expected to formulate their own observations, evaluate their own data and reach their own conclusions. This discovery approach encourages curiosity and observation, inquisitiveness and speculation. Careful thinking habits and the ability to search for cause and effect and to make cautious conclusions are developed. Pupils must critically evaluate evidence and be critical of unsupported statements, while remaining open-minded and tolerant of the opinions of others (1967, page 1.).

Implicit in this statement is the assumption that there are various levels of thinking and that students exposed to the content and method of the "new" science courses would develop their critical thinking abilities to levels beyond that of "mere memorization and recall." The expectation seems to be that with increased exposure to science programs of a specific nature, habits of critical thinking would be fostered. Implicit also in this statement is the assumption that critical thinking ability among students is identifiable and measurable.

According to Watson and Glaser (1964), critical thinking is a composite of attitudes, knowledge and skills. Consequently, for them, in educational and psychological terms, the student who is high in critical thinking ability is characterized by:

1. Attitudes of enquiry that involve the ability to recognize the existence of problems and the acceptance of the general need for evidence in support of what is asserted to be true.

2. Knowledge of the nature of valid inferences, abstractions and generalizations in which the weight of accuracy of different kinds of evidence are logically determined.

3. Skills in employing the above attitudes and knowledge.

For the purposes of this study, the operational definition of critical thinking, proposed by Watson and Glaser, was adopted and critical thinking was measured by the Watson-Glaser Critical Thinking Appraisal. In this test, critical thinking is defined operationally in terms of the combined score from five sub-tests, namely: inference, recognition of assumptions, deduction, interpretation and evaluation of arguments. Implicit in the use of this test is the assumption that there exists within this list of abilities designated as critical thinking sufficient overlap to warrant the expectation that the subtest scores will have a very high correlation with the total score. A more complete description of critical thinking, as a valid educational outcome, and of the Watson-Glaser Critical Thinking Appraisal, as a measure of critical thinking, is to be found in Chapters II and III respectively.

On the basis of the literature reviewed, it would appear that students in the upper grades could be expected to achieve higher scores in critical thinking than those in the lower grades. For the purposes of the study then, it was necessary to ask what independent variables might be important in any gain observed in critical thinking over the time span of the study. Could such gain be attributable to age or sex? Would older students gain more or less than younger ones? Would sex at this stage of development have any effect? Would students of initially higher levels of critical thinking ability tend to make greater gains? Would the number of science courses completed affect the size of the gain? Would there be any effect on the gain that could be attributed to the type of course taken, either university entrance or general science? And finally, within the science types, would there be any differences in gain that could be attributable to specific subject matter content?

It follows that the study should be limited to seeking answers to the following questions:

1. Is there an increase in critical thinking ability of senior high school students between grades ten and twelve?

2. Is there a difference in the mean gain in critical thinking of the sexes?

3. Do students high in critical thinking ability tend to select a greater number of specific science courses?

4. Is there a difference in the mean gain in critical thinking ability of science and non-science students?

5. Among the science students, is there a difference in the mean gain in critical thinking of university entrance and general course students?

6. Among university entrance and general course students, is there a difference in the mean gain in critical thinking of students who take and complete more science courses than of those who do not?

It should be noted that this study did not seek to compare the critical thinking ability of students pre-matched on the bases of age, sex and grade who subsequently were enrolled in specific science or non-science programs. Nor did the study seek to compare the pre- and post-performances of students, in terms of course grades, who had been exposed to a methodology that was designed to foster critical thinking as opposed to those who had not. Instead, the study sought to identify differences, if any, in the critical thinking abilities of students who, by reason of parental or other guidance or by reason of individual choice, have pre-selected themselves into the categories labelled science/non-science.

Definition of Terms

For the purposes of this study, students at the time of post-testing were classified as (1) a science major if they had successfully completed four or more science courses, (2) a science minor if they had completed successfully not more than three science courses, and (3) a non-science student if they had completed fewer than two science courses. Within these categories, a further distinction was made between university entrance course students and general course students. A general course student did not take any university entrance courses.

In making these distinctions, it is to be noted that (1) a general science course is a course designated as an "01" course, e.g., biology 201, 301; physical science 201, 301; and (2) a university entrance science course is a course designated as a "00" course, e.g., biology 200, 300; chemistry 200, 300. Further, the numbers 200 and 201 refer to courses normally taken in the grade eleven program and the numbers 300 and 301 refer to courses normally taken in grade twelve.

Justification of the Study

A review of the literature suggested that there was room for a study, the findings of which could prove useful in future curriculum planning. Such a study could answer to some degree the question of the relationship between the selection of science courses as opposed to non-science courses and critical thinking ability. It could also seek to discover whether the ability to think critically is a factor in determining which science courses, if any, are selected. And further, it could help to determine the relationship between growth in critical thinking and exposure to science courses.

Plan of the Thesis

In this chapter, a statement of the problem of this study has been made. Chapter II contains a summary of the literature relevant to this problem. Chapter III contains an outline of the study conducted. The results of the study are presented in Chapter IV. The summary and conclusions to be drawn from the results follow in Chapter V.

CHAPTER II

REVIEW OF THE RELATED LITERATURE

When one examines the available literature, one is struck by the fact that, although critical thinking is often cited as a desirable goal of education, exactly what is meant by the term or how one can judge (i.e., measure) critical thinking is rarely discussed. There are, then, two purposes to which a review of the literature might be addressed: (1) to attempt to delineate more clearly what is meant by the term <u>critical thinking per se</u> and (2) to examine the results of the research into the achievement of critical thinking as a valid goal of science education. Both of these objectives are pursued in this review.

The Literature Pertaining to the Concept of Critical Thinking

A review of the available literature reveals that psychologists and educators do not always agree on the precise nature of critical thinking. Consequently, the definition given for critical thinking is found to vary from writer to writer. In this section, an attempt is made to identify the common elements among the various definitions and so to arrive, for the purposes of this study, at a satisfactory operational definition of the term <u>critical thinking</u>.

Critical thinking has been defined by Burmester (1952) as encompassing most if not all of the following abilities: (1) to recognize a problem, (2) to delimit a problem, (3) to recognize and accumulate facts related to a problem, (4) to recognize and formulate an hypothesis, (5) to plan an experiment to test an hypothesis, (6) to carry out an experiment to test an hypothesis, (7) to interpret data collected, and (8) to generalize the conclusion to a new situation.

The similarity of this list of abilities to those that are required for the successful application of the scientific method is obvious. But, if critical thinking is the ability (or abilities) required to define a problem, to recognize assumptions, to formulate relevant hypotheses, to select pertinent information, and to draw conclusions validly (Dressel and Mayhew, 1954), how is this ability to be distinguished from general intelligence?

Utilizing tasks that are not commonly used in the construction of intelligence tests, Watson and Glaser (1964) reported substantial correlations between the scores on their critical thinking appraisal and the scores on various verbal intelligence test scores (for complete details, see Chapter III). Their findings were, later, substantiated by Haas (Skinner, 1971).

Coefficients ranging from .55 to .75 with a median $\underline{r}$ of .68, are evidence of a substantial relationship between critical thinking and mental ability scores as measured by conventional intelligence tests. Further light on the relationship between critical thinking and general intelligence may be shed by the results of factor analytic studies and, in particular, by those undertaken by J. P. Guilford.

In his Structure of Intellect, Guilford (1963) recognizes at least one hundred and twenty factors which he classifies according to three dimensions: cognitive, productive and evaluative. According to him, whenever a process or problem is to be solved and the constructs of the situation are required, the cognitive factors are at play. Then, the production factors are brought into action until a solution is achieved by either a divergent or a convergent thinking process. And finally, the evaluative factors are utilized to determine the suitability or the effectiveness of the thinking process.

In terms of Guilford's model, Ennis (1969) believes that critical thinking corresponds closely to convergent thinking and evaluation. But, others support the concept of critical thinking as problem solving which is a multi-factor process (Johnson, 1962). Allan and Rott (Madison, 1964) conclude that critical thinking should be regarded as a "pluralistic act" including an evaluative process.

It appears, then, that a high level of general intelligence may be prerequisite to a high level of critical thinking but it is not clear in any of these definitions whether critical thinking is to be identified with reflective thought in the Dewey sense (Dewey, 1923), with formal reasoning in the Piagetian sense (Inhelder and Piaget, 1964) or with problem solving in the Gestalt sense (Wertheimer, 1945).

Burton, Kimball and Wing (1971) define critical thinking as "the critical reflective search for valid conclusions which solve our problems, resolve our doubts, and enable us to choose between conflicting statements of doctrine or policy." Their concept of critical thinking, as delineated by a critical thinking scale (1971, pp. 450-451) is supported by Usery (Madison, 1964) who considers critical thinking to be "an act of searching for the clearest ideas about a subject derived from the facts, points of view, observations, and other elements."

Smith (Madison, 1964), on the other hand, argues that critical thinking is characterized by "good unemotional judgement that results from an analysis of the situation or of the materials." He is supported in this view by Lien (1967) who concludes that, in the weighing of evidence and in the answering of the question: What are the logical results?, the underlying elements in critical thinking involve the distinguishing of facts from opinion, the drawing of inferences, and the drawing of valid conclusions.

Ennis (Troost, 1971) also supports the idea that critical thinking involves reasoning. Ennis emphasizes that critical thinking is characterized by "the correct assessment of statements" and lists the following situational aspects:

1. Grasping the meaning of a statement;
2. Judging whether or not there is ambiguity in a line of reasoning;
3. Judging whether or not certain statements contradict each other;
4. Judging whether or not a situation is actually the application of a specified principle;
5. Judging whether or not a statement is actually specific enough;
6. Judging whether or not an observation is reliable;
7. Judging whether or not a conclusion necessarily follows;
8. Judging whether or not an inductive conclusion is warranted;
9. Judging whether or not the problem has been identified;
10. Judging whether or not a statement is an assumption;
11. Judging whether or not a definition is adequate;
12. Judging whether or not a statement made by an alleged authority is acceptable.

It appears then that critical thinking can be regarded as "a psychological or mental process in which a pupil draws from his knowledge background plus an ability to use logical reasoning in an effort to avoid common errors in judgement" (Brown and Brown 1971). But, there remains a problem still. If critical thinking is to be measured, it must be defined operationally in terms of a number of sub-steps through which an investigator must carry his search in order to bring it to fruition.

Typical of this approach is that of Burton, Kimbal and Wing (1964) who provide in classical form, the following summary of this process:

1. Recognition and definition of a problem.
2. Hypothesis formulation of possible solutions.
3. Inquiry or search which involves procedures of (a) experimentation, (b) collection of data, and (c) reasoning by induction, deduction and analogy.
4. Decision and acceptance of an hypothesis.
5. Testing and the use of the accepted conclusion.

Ennis (1969), on the other hand, provides a more detailed and more useful description of the application of a subject's critical thinking ability in terms of specific problem situations. According to him, a subject having a high level of critical thinking ability can:

1. Give illustrations of his own (Can you give another example of this?)
2. Relate facts to past experiences (What do you already know about this?)
3. Apply facts which relate to his own life (How does this relate to you now?)

4. Draw inferences (What does this mean?)
5. Apply facts to his local community (How does this relate to your town or city?)
6. Weigh the evidence presented in two or more accounts of the same subject (What do you believe?)
7. Read between the lines of print (What is actually being said?)
8. Draw conclusions from facts (What is the logical conclusion?)
9. Distinguish facts from opinion (What is the evidence?)
10. Select pertinent facts (What are the main ideas?)
11. Reason from cause to effect (What is the logical result?)

It is clear that Ennis concedes that critical thinking involves a number of heuristic skills (inference, recognition of assumptions, deduction, induction, interpretation, evaluation). Critical thinking is not a singular skill but a generic term describing both a process and an ability (D'Angelo 1964). It is composed of attitudes, knowledge and skills and these skills are based on sound judgment and allied with problem solving.

The operational definition used by Watson and Glaser in the development of their critical thinking appraisal comes closest to this concept. Their test was designed to measure: (1) attitudes of inquiry that involve an ability to recognize the existence of a problem and an acceptance of the general need to support what is asserted to be true with evidence, (2) knowledge of the nature of valid inferences, abstractions and generalizations in which the weight of accuracy of different kinds of evidence are logically determined and (3) skill in employing or applying the above attitudes and knowledge.

Literature on the Achievement of Critical Thinking as a Goal of Science Education

Critical thinking is cited as an important educational objective in many programs of studies. For example, the Manitoba High School Program of Studies Grades 9-12 (1975-1976) states, "For the majority, the prime objectives of public school education reflect the traditional concerns of the transmission of knowledge, the acquisition of critical inquiring habits of thought and adequate preparation either to obtain immediate employment or some form of post secondary education or training."

Similar references to the importance of critical thinking are found in the teachers' manuals for specific science courses such as B.S.C.S. biology in which the philosophy of the course is that students should discover science rather than learn only the facts of science. The development of critical thinking is proposed as a goal deserving high priority and it is suggested that there is no way that schools can better prepare students for life than by helping them develop their powers of reasoning.

Science educators also support critical thinking as a desirable goal of education. Voss and Brown (1968) state that trends in testing as in the teaching of science have been away from an emphasis on the recall of factual material and toward an emphasis on critical thinking, the understanding of the relationships between different concepts learned and on ability to apply knowledge in new situations. They suggest further that critical thinking is a prerequisite for a student to carry on independent inquiry in which he can use his knowledge of science and his skills of the processes of science to ask the proper questions about science. In turn, these skills induce the explanation of the hypotheses that can lead him to further inquiry.

Schwab (1962) provides an excellent opportunity for development of inquiry in his forty-four "invitations to enquiry". These "invitations" are examinations of research reports complete with appropriate questions for class and student discussion. The use of these invitations provides guided assistance to the teacher and student in the development of critical thinking.

Ausubel (1965) questions the value of inquiry in science instruction as to its philosophical and psychological foundations for the majority of the large scale curriculum developments in the last ten years. Instead he stresses the idea of structure of material as the most fundamental goal of science teaching. This theory has drawn much support representing the symbolic use of word signs for teaching the concepts of science because this provides an orderly organization of material.

Gagne (1963) proposes that students move through various levels of development in order to become a participant of independent inquiry. He further believes that students may leave science at various levels which include the methods of acquiring knowledge with some practice in inquiry. The highest level is one in which the student assumes the role of a truly independent investigator, capable of looking at problems objectively and to have ideas and judge them critically.

Glass (Schwab 1962) adopts the point of view that the primary aim in science teaching is to lead science students to understand scientific methods of investigations and to appreciate the spirit and the outlook of a scientist. He further states that it is more important that students have an opportunity to move step by step through acquiring the necessary skills to experience the joy of discovery. Students must learn to ask the right questions, to form testable hypotheses, to draw valid conclusions from their data and determine the significance of their findings.

It would appear that critical thinking is considered to be an important educational goal both by departments of education and science educators.

Critical Thinking and High School Science Courses

While science educators concede that critical thinking is a desirable educational goal, what evidence is there from research that this goal is being attained? What is the relationship of critical thinking to science exposure? Is critical thinking teachable?

A survey of the literature reveals that there has been limited research of critical thinking related to courses taken in science.

In a study by George (1965) in which a comparison of B.S.C.S. biology students in one group used the blue version as course material scored significantly higher on the Watson-Glaser Critical Thinking Appraisal than those in another group in traditional biology when IQ was controlled. No differences were found between the green and yellow B.S.C.S. versions and those in traditional biology courses.

Two methods of teaching high school biology were compared in a study by Kastrinos. One group designated, "the text book recitation group" was compared to another group, "the principles critical group." This latter group scored higher in critical thinking indicating that methodology can influence critical thinking in the teaching of science courses.

A study completed by Charen (1970) was designed to determine whether laboratory methods stimulate critical thinking in chemistry. Different classes of high school students were exposed to two laboratory techniques in which one group used materials and methods intended to promote critical thinking while the other used standard procedures. Mastery of the attributes of critical thinking included ability associated with the nature of proof, ability to interpret data, ability to recognize and make assumptions, ability to test and evaluate evidence, in fact most of the sub-categories proposed by Watson and Glaser.

The tests of critical thinking in chemistry showed no significant differences in favor of either laboratory approach. An analysis of the scores of the Watson-Glaser Critical Thinking Appraisal demonstrated a significant difference in favor of the traditional classes. Some of the reasons suggested for the failure to improve critical thinking during the laboratory treatment are stated in the study. These include too short a timespan to permit the effects of the new methods to be utilized by the students and also that the short exposure to the open-end technique may have produced the phenomenon of interference.

In a study by Henkel (1967) the critical thinking of college entrance students studying P.S.S.C. physics was compared to students enrolled in traditional physics courses. Measurement of critical thinking was by the W.G.C.T.A. The conclusion reached in the study was that P.S.S.C. students had significant gains in critical thinking with greater growth shown for students with previous science courses.

Rickert (1967) designed a special course in an attempt to improve the critical thinking of students. It provided an opportunity for students to analyze problems, examine assumptions, collect and organize data and test hypotheses in the physical sciences. He compared the results of students receiving the experimental treatment to students in two other physical science courses using the A.C.E. test of critical thinking and STEP tests. His conclusion was that the experimental course students were superior in developing "ability to think critically" for all ability groups. This indicates that critical thinking can be improved by courses of this type.

Other studies have been based on the teacher's role in fostering critical thinking. One such study carried out by Kleinman showed that students with teachers who asked more critical thinking questions did significantly better on a test-retest of understanding science. This study suggests that teachers can have an effect on the critical thinking ability of their students through the use of critical thinking questions.

George and Dietz (1968) carried out a study to assess the effect of teacher's critical thinking on students. The Watson-Glaser Critical Thinking Appraisal was used in evaluation. The results were that students exposed to teachers ranking high in critical thinking ability achieved the highest mean score on the posttest while students exposed to teachers low in critical thinking ability achieved the lowest adjusted mean scores on the posttest. In general, grade level could not be considered significant in influencing critical thinking, a result that does not agree with those of Watson and Glaser.

Tate and Stanier (1964) analyzed the performance of good and poor problem solvers on tests of critical thinking and practical judgment. The subjects were high school students whose scores on a composite measure of problem solving ability including mathematical and quantitative reasoning problems deviated markedly from a regression line of problem solving on IQ. On the critical thinking tests the poor problem solvers tended to avoid the judgment "not enough facts" and to make unqualified true or false judgments.

If course exposure, methods of instruction and the critical thinking ability of the classroom teacher can influence the critical thinking ability of students, then one may ask what is the effect on critical thinking of contemporary high school courses? The following chapter is directed towards answering this question.

CHAPTER III

METHODOLOGY AND PROCEDURES

This chapter is a definitive description of the important elements of the study. The description includes: the statement of the problem; some methodological questions; the sample and its composition; the sampling procedures; the test instrument; the administration procedures in the study and the plan for the analysis of the data.

Statement of the Problem

This study sought to establish whether differential exposure to science courses was related to gain scores on the dimension of critical thinking. This aim raised a series of methodological questions such as:

1. How was exposure to science to be measured?
2. How was critical thinking to be measured?
3. How was gain in critical thinking to be measured?

Exposure to Science

For the purpose of the study exposure to science courses was defined as the number of science courses completed in grades eleven and twelve. It was recognized that students during this period could take three different types of programs: all university entrance, all general courses, and a combination of these. This fact suggested three types of science exposure: university entrance, mixed, and general science.

Definition of Critical Thinking

A problem in the study was to measure the growth of critical thinking ability in high school students as a function of science courses completed in grades eleven and twelve. For this purpose the scores on the Watson-Glaser Critical Thinking Appraisal Form YM and ZM were used. That is, for the purpose of the study the operational definition utilized by Watson and Glaser in the development of their test was used.

Gain in Critical Thinking

In the standardization of their test, Watson and Glaser reported different means and standard deviations for the different forms. Hence a problem existed in the use of Form YM as pretest and Form ZM as posttest. The forms could not be considered equivalent. Consequently, for the purposes stated a student's gain score was defined as the difference between the pretest and posttest z-scores on the Watson-Glaser Critical Thinking Appraisal.

The Sample

The sample, for the purposes of this study, was composed of two hundred twenty high school students enrolled for the 1973-74 school year in grades ten, eleven and twelve in a senior high school in the province of Manitoba.

At the time of pretesting on the Watson-Glaser Critical Thinking Appraisal, the sample included all of the students in the school which was judged representative of the school population, a population similar to that used by Watson and Glaser who omitted extremely large cities from their sample because they did not find them representative of the entire population. The sample was considered representative for the following reasons: First, both rural and urban students were included. Second, the courses of instruction followed by students included academic, industrial, business education. Further evidence that the sample is representative of the Watson-Glaser population is shown in Table 1 in which the means of the Watson-Glaser Critical Thinking Appraisal scores from the sample used in the study and from the samples used by Watson and Glaser are compared.

TABLE 1

WATSON-GLASER MEAN SCORES BY GRADE
FOR MANITOBA HIGH SCHOOL AND
FOR WATSON-GLASER SAMPLE
(cf. Table 5d Watson and Glaser, 1964)

	MANITOBA SAMPLE			WATSON-GLASER SAMPLE		
	Number of Students	Mean Score	St. Dev.	Number of Students	Mean Score	St. Dev.
Grade X	78	61.62	10.1	2,947	61.7	11.0
Grade XI	67	64.98	10.6	2,406	64.4	11.0
Grade XII	75	66.18	9.04	1,800	65.6	10.9
Grades X, XI, and XII	220	64.2	10.1	7.153	63.58	

Description of the Instrument

The Watson-Glaser Critical Thinking Appraisal was developed over a period of twenty-five years of research, on the measurement of critical thinking abilities. The early tests were developed by Watson for his study, "The Measurement of Fairmindedness". Later in 1937, Glaser revised the tests for use in "An Experiment in the Development of Critical Thinking". After this study was completed, a number of successive analyses and refinements were made with consideration given to the theoretical concepts of critical thinking as well as to the more practical questions posed by people who used the tests.

In preparation of the revised edition of the "Critical Thinking Appraisal", the items of earlier tests were reviewed and new items were written to replace those which had been questioned by those who used and reviewed the test. The performance of these items as well as those retained from earlier editions was studied in experimental programs at high schools as well as at college levels and also with various industrial personnel groups. From these studies, item-analysis statistics and comments from various well qualified persons served as a basis for refining further test items.

Consideration was given to the amount of time required by students to complete all test items and to the failure of students to complete the various subtests. This made it possible for the establishment of more realistic time limits for the revised edition. The directions for the subtests were rewritten in view of recommendations made in reviews of the tests and by others who critically evaluated the tests.

The result is a battery of those tests and test items which were found to be most functional and significant and which appeared to measure critical thinking.

Once the forms YM and ZM of the Critical Thinking Appraisal had been revised according to criteria set out by the authors and publisher, the tests were administered to representative samples of well defined populations of high school students and college freshmen.

The Watson-Glaser Critical Thinking Appraisal, the Quick Word test, the Otis Quick-Scoring Mental Ability Tests, and the Gamma Tests were administered in fourteen school systems in thirteen states in 1963. By alternating the forms of the test, one half of the random sample took YM while the other half took ZM.

Those high schools permitted to participate in the standardization program had to be a regular public institution in a community of 10,000 to 75,000 with at least 100 students at each of the four grade levels nine to twelve. The sample was chosen in this way to avoid any possible bias associated with schools that were extremely small, and with highly specialized high schools found in some very large systems. Selective influences were further reduced by the requirement that all students be tested in the participating schools.

The results of the tests from the 20,312 students tested were then analyzed and used to provide comparative normative information for each form and grade level. Correlations between raw scores on the Watson-Glaser Critical Thinking Appraisal and various Verbal Intelligence measures are reported for several schools and adult groups in Tables A-1 and A-2, Appendix A.

The college standardization program was conducted in 1963 with 5,297 freshmen at fifteen four-year liberal arts colleges in eleven states (Madison, John P., 1964). The enrolments ranged from 200 to 1,000 students in the institutions tested with the Critical Thinking Appraisal Form YM and the Otis Gamma Test to provide the normative data as shown in Table A-3 in Appendix A.

Houle (1943) and Morse (1957) support the belief that the items in the Watson-Glaser Critical Thinking Appraisal represent an adequate and reliable sample of the five abilities and further that the total score of the test by students represents a valid estimate of the proficiency of individuals with respect to these aspects of critical thinking.

Permission to Conduct the Study

Permission to carry out the study was obtained from the superintendent of the public school division, from the principal of the high school and from the teachers of the various classes in which the tests were conducted. Permission was also granted to obtain subject grades from the school records at the end of the school year in June, 1973.

Data Collection Procedures

The Watson-Glaser Critical Thinking Appraisal Form YM was administered to two hundred twenty high school students in grade ten, in May of the school year 1973. Since the sample was found to be representative, only one school was used. This procedure was utilized to eliminate as many variables, such as teaching techniques, courses of instruction, philosophy, and attendance policy, as possible. May was chosen to avoid the anxieties sometimes associated with the year end testing programs.

The Watson-Glaser Critical Thinking Appraisal Form ZM, was administered to sixty-two students from the original (1973) sample at the completion of grade twelve in June 1976. Because of population mobility and natural attrition the original sample was unavoidably reduced.

The tests were written in their regular classes as they occurred on their regular timetable. This schedule facilitated the sorting, and marking of the tests in their respective science courses.

The science classes tested included: science 100, 101; biology 200, 201, 300, 301; chemistry 200, 300; physics 200, 300; physical science 201, 301.

The students marked their choices of multiple choice questions on computer cards. This brought uniformity to the writing of the test and made it possible to make use of the computer terminal in the school to score the test and to obtain some statistical data, removing as much as possible the risk of human error.

Directions were dictated to the students using a standard set of instructions both in completion of student's name, course, age, etc., as well as for the marking of the test items of their choice on the cards. See Appendix B.

The directions for writing the test were provided to the students by the author. This standard set of instructions taken from the test manual was included with the class package. The instructions included the procedure to be followed if a student should make an incorrect choice and wished to alter it. Also, students were cautioned regarding the use of time so that all questions could be attempted. Further details of the instructions may be found in Appendix B.

Plan for the Analysis of the Data

At the conclusion of the data collecting phase of the study, the following data were available for analysis:

1. Scores from Watson-Glaser Critical Thinking Appraisal Form YM and ZM.
2. School records of subjects taken by students during grades ten, eleven and twelve.

The statistical procedures of (1) analysis of variance and (2) correlation and regression analysis, were deemed appropriate to test the following null hypotheses at the .05 level of significance.

Hypothesis A: There is no difference between the critical thinking by students in grade ten and grade twelve.

Hypothesis B: There is no difference in the mean gain in critical thinking of students classified by age and sex.

Hypothesis C: There is zero correlation between the critical thinking of students obtained in grade ten and the number of science courses they took in their science program.

Hypothesis D: There is no difference in the mean gain in critical thinking of science and non-science students.

Hypothesis E: There is no difference in the mean gain in critical thinking of university entrance and general course science students.

In order to test this hypothesis the following subhypotheses were made:

Subhypothesis E-1: There is no difference in the mean gain in critical thinking by students in university entrance physics and general course physical science.

Subhypothesis E-2: There is no difference in the mean gain in critical thinking by students in university entrance biology and general course biology.

Subhypothesis E-3: There is no difference in the mean gain in critical thinking by students in university entrance chemistry and general course physical science.

Subhypothesis E-4: There is no difference in the mean gain in critical thinking by students in university entrance mathematics and general course mathematics.

Hypothesis F: There is no difference in the mean gain of critical thinking among university entrance students and general course students as a function of the number of science courses taken.

CHAPTER IV

THE RESULTS OF THE STUDY

In the previous chapter a study was outlined in which an attempt was made to determine if the gain in critical thinking ability of senior high school students is related to their science education in terms of the number and type of science courses completed. In this chapter the data that were collected are examined with respect to a number of null hypotheses posed.

Hypothesis A: There is no difference in the mean scores in critical thinking of students in grade ten and grade twelve.

In order to test this hypothesis it was necessary, because Form YM (used for pretest) and Form ZM (used for the posttest) were not equivalent tests, to transcribe the scores on the pretests (given in grade ten) and the posttest (given in grade twelve) into standard scores. The resulting frequency distributions are shown in Table 2.

H_o: $\bar{d}_z = 0$ was tested by means of the t-test for the difference between means for correlated samples. The alternative hypotheses H_a: $\bar{d}_z \neq 0$ indicated a two tailed test. The critical value of t with 61 degrees of freedom is 2.66 at .01 level of significance. The observed t was found to be 6.623. H_o was therefore rejected.

The statistics from which the observed t can be calculated are found in Table C-1, Appendix C.

TABLE 2

Z-SCORE DISTRIBUTIONS FOR WATSON-GLASER PRETEST, POSTTEST, AND GAIN SCORES

Z SCORE	FORM YM	FORM ZM	GAIN
2.00 - 2.20		1	
1.80 - 2.00	1	1	
1.60 - 1.80	1	5	2
1.40 - 1.60	1	3	2
1.20 - 1.40	0	3	1
1.00 - 1.20	2	2	1
.800 - 1.00	3	2	7
.600 - .800	7	7	9
.400 - .600	7	10	12
.200 - .400	8	4	10
.000 - .200	1	3	14
.000 - -.200	2	4	1
-.200 - -.400	4	1	0
-.400 - -.600	3	2	0
-.600 - -.800	4	5	1
-.800 - -1.00	5	4	0
-1.00 - -1.20	7	2	1
-1.20 - -1.40	1	2	1
-1.40 - -1.60	1	0	
-1.60 - -1.80	2	0	
-1.80 - -2.00	1	1	
-2.00 - -2.20	1		
	N = 62	N = 62	N = 62

Hypothesis B: There is no difference in the mean gain in critical thinking of students classified by sex and age.

The data to test Hypothesis B by analysis of variance are found in Table 3. In this and subsequent tables, N = d.f. + 1 (i.e., N = 62).

TABLE 3

WATSON-GLASER CRITICAL THINKING GAIN SCORES
BY SEX AND AGE

SOURCE OF VARIANCE	SUM OF SQUARES	DF	MEAN SQUARES	F	P
Main Effects	1.61	4	0.40	1.36	0.26
Sex	0.11	1	0.11	0.38	0.54
Age	1.49	3	0.50	1.68	0.18
2-Way Interactions					
Sex x Age	0.04	2	0.02	0.06	0.94
Explained	1.64	6	0.27	0.93	0.48
Residual	16.25	55	0.30		
Total	17.90	61	0.29		

There were no significant interactions. Neither were there any significant main effects. For sex, $P(F = 1.38) \geq .26$; for age, $P(F = 1.68 \geq .18)$. Further data relating to the testing of this hypothesis are presented in Table C-2, Appendix C. It should be noted that on the basis of the data collected, Hypothesis B could not be rejected.

Hypothesis C: There is zero correlation between the critical thinking of students obtained in grade ten and the number of science courses they took in their science program.

To test Hypothesis C a number of Pearson correlation coefficients were computed as shown in Table 4. The Pearson r between the Watson-Glaser pretest z-scores (obtained in grade ten) and the number of science courses subsequently completed was found to be .32.

The null hypothesis H_o: $r_{xy} = 0$ was tested by means of a two-tailed t test for correlated samples. The significance of the observed r was .006. The null hypothesis was rejected.

The Pearson r between Watson-Glaser pretest z-scores and the number of university entrance science courses was found to be .44. The significance of the observed r was .001. In terms of the number of university entrance science courses subsequently completed, the null hypothesis was rejected.

The Pearson r between Watson-Glaser pretest z-scores and the number of general science courses completed was found to be -.32. The significance of the observed r was .006. In terms of the number of general science courses subsequently completed the null hypothesis was rejected.

As can be seen from the data in Table 4 a similar result was found for the Watson-Glaser posttest z-scores and the Watson-Glaser gain z-scores.

TABLE 4

PEARSON CORRELATION COEFFICIENTS BETWEEN WATSON-GLASER CRITICAL THINKING Z-SCORES AND NUMBER OF SCIENCE COURSES

	U.E.	GENERAL	COMBINED
PRETEST	0.44	-0.31	0.32
P(r)	0.001	0.006	0.006
N	62	62	62
POSTTEST	0.64	-0.44	0.48
P(r)	0.001	0.001	0.001
N	62	62	62
GAIN	0.35	-0.21	0.27
P(r)	0.003	0.049	0.015
N	62	62	62

Hypothesis D: There is no difference in the mean gain in critical thinking of science and non-science students.

In order to test this hypothesis it was necessary to define science and non-science students. Science students were separated into two categories or classes. A science major was defined as a student who completed four or more science courses and a science minor as one who completed two or three science courses. A non-science student was defined as a student who completed zero or one science course.

The data to test Hypothesis D by analysis of variance are found in Table 5.

TABLE 5

WATSON-GLASER CRITICAL THINKING GAIN SCORES
BY AGE, SEX AND CLASS

SOURCE OF VARIANCE	SUM OF SQUARES	DF	MEAN SQUARE	F	P
Main Effects	3.23	6	0.54	2.20	0.06
Age	1.06	3	0.35	1.45	0.24
Sex	0.25	1	0.25	1.01	0.32
Class	1.62	2	0.81	3.31	0.05
2-Way Interactions	3.19	8	0.40	1.63	0.14
Age x Sex	0.20	2	0.10	0.41	0.67
Age x Class	1.99	4	0.50	2.04	0.10
Sex x Class	0.51	2	0.26	1.05	0.36
Explained	6.41	14	0.46	1.88	0.06
Residual	11.48	47	0.24		
Total	17.90	61	0.29		

Since there were no significant interactions, the main effects could be tested. Class was shown to be significant [$P(F = 3.31) \leq .05$]. Age was not a significant variable [$P(F = 1.44) \geq .24$] and neither was sex [$P(F = 1.01) \leq .32$].

Other data relating to testing the means in this hypothesis are found in Table C-3 Appendix C. In terms of the number of science courses completed the null hypothesis was rejected.

<u>Hypothesis</u> E: There is no difference in the mean gain in critical thinking of university entrance and general course students.

In order to test this hypothesis the following subhypotheses were made:

Subhypothesis E-1: There is no difference in the mean gain in critical thinking by students in university entrance physics and general course physical science.

In testing the subhypotheses E-1, E-2, E-3, and E-4 the students were defined as university entrance students if they completed two university entrance science courses and as general course students if they 1) completed zero or one university entrance science course taken concurrently with a university entrance science course or 2) if they completed two general science courses.

The data used to test subhypothesis E-1 are presented in Table 6.

TABLE 6

WATSON-GLASER CRITICAL THINKING GAIN SCORES FOR PHYSICS STUDENTS BY SEX AND TYPE

SOURCE OF VARIANCE	SUM OF SQUARES	DF	MEAN SQUARE	F	P
Main Effects	0.79	2	0.39	1.29	0.30
Sex	0.55	1	0.55	1.82	0.19
Type	0.36	1	0.36	1.18	0.29
2-Way Interactions					
Sex x Type	0.01	1	0.01	0.04	0.84
Explained	0.80	3	0.27	0.88	0.47
Residual	6.09	20	0.30		
Total	6.89	23	0.30		

There are no significant interactions. Neither are there any significant main effects; sex [P(F = 1.82) ≥ .19]; and type [P(F = 1.18) ≥ .29]. On the basis of this data subhypothesis E-1 was not rejected. Further data regarding the means used in testing subhypothesis E-1 are found in Table C-4, Appendix C.

Subhypothesis E-2: There is no difference in the mean gain in critical thinking by students in university entrance biology and general course biology.

The data used to test subhypothesis E-2 by analysis of variance are found in Table 7.

TABLE 7

WATSON-GLASER CRITICAL THINKING GAIN SCORES FOR BIOLOGY STUDENTS BY SEX AND TYPE

SOURCE OF VARIANCE	SUM OF SQUARES	DF	MEAN SQUARE	F	P
Main Effects	0.82	2	0.41	1.11	0.34
Sex	0.27	1	0.27	0.73	0.40
Type	0.27	1	0.27	0.73	0.40
2-Way Interactions					
Sex x Type	0.04	1	0.04	0.11	0.74
Explained	0.86	3	0.29	0.78	0.51
Residual	13.62	37	0.37		
Total	14.48	40	0.36		

There were no significant interactions. Nor were there any significant main effects; sex [P(F = .73) ≥ .40], type, U.E. biology or general biology, [P(F = .73) ≥ .40]. In terms of biology courses completed subhypothesis E-2 was not rejected. Further data regarding the means tested in this hypothesis are found in Table C-5, Appendix C.

Subhypothesis E-3: There is no difference in the mean gain in critical thinking by students in university entrance chemistry and physical science.

The data used to test subhypothesis E-3 by analysis of variance are found in Table 8.

TABLE 8

WATSON-GLASER CRITICAL THINKING GAIN SCORES FOR CHEMISTRY STUDENTS BY SEX AND TYPE

SOURCE OF VARIANCE	SUM OF SQUARES	DF	MEAN SQUARE	F	P
Main Effects	1.68	2	0.84	5.16	0.02
Sex	0.27	1	0.27	1.64	0.21
Type	1.28	1	1.28	7.90	0.01
2-Way Interactions					
Sex x Type	0.00	1	0.00	0.01	0.91
Explained	1.68	3	0.56	3.45	0.04
Residual	3.08	19	0.16		
Total	4.76	22	0.22		

There were no significant interactions. In testing main effects, sex was found to be not significant [$P(F = 1.64) \geq .21$] but type was found to be significant [$P(F = 7.90) \leq .05$]. In terms of chemistry courses completed, subhypothesis E-3 was rejected. Further data regarding the means tested in this hypothesis are found in Table C-6, Appendix C.

Subhypothesis E-4: There is no difference in the mean gain in critical thinking by students in university entrance mathematics and general course mathematics.

The data used to test subhypothesis E-4 by analysis of variance is found in Table 9.

TABLE 9

WATSON-GLASER CRITICAL THINKING GAIN SCORES FOR MATHEMATICS STUDENTS BY SEX AND TYPE

SOURCE OF VARIANCE	SUM OF SQUARES	DF	MEAN SQUARE	F	P
Main Effects	4.16	2	2.08	8.71	0.00
Sex	0.48	1	0.48	2.02	0.16
Type	3.78	1	3.78	15.81	0.00
2-Way Interactions					
Sex x Type	0.00	1	0.00	0.00	0.97
Explained	4.16	3	1.39	5.80	0.00
Residual	11.47	48	0.24		
Total	15.64	51	0.31		

There were no significant interactions. Among main effects, sex was not significant [P(F = 2.02) $\geq$ .16]. Type (i.e., the number of mathematics courses completed) was found to be significant [P(F = 15.8) $\leq$ .05]. Data regarding the means tested in this hypothesis are found in Table C-7 of Appendix C. The greatest gains were made by those students who completed the greater number of university entrance mathematics courses.

Hypothesis F: There is no difference in the mean gain in critical thinking among university entrance and general course science students as a function of the number of science courses taken.

In order to test hypothesis F, students were classified according to type and class. The variable type had three levels: (1) university entrance (students had taken two university entrance science courses) (2) mixed (students had taken only one university entrance and one general science course) and (3) general (students had taken only general science courses). It must be noted that, in this analysis, no distinction was made according to subject and mathematics was not included as a science course.

The variable class had three levels: (1) science major (2) science minor and (3) non-science. See page six, Chapter I.

The data to test hypothesis F are presented in Table 10. There were no significant interactions. Among main effects, type was not significant [P(F = 1.20) $\geq$ .31]. However, class had a significant effect [P(F = 3.21) $\geq$ .05]. Further data with respect to the means tested in hypothesis F are found in Table C-8 in Appendix C. The greatest gains in critical thinking were made by the science majors.

TABLE 10

WATSON-GLASER CRITICAL THINKING GAIN SCORES FOR STUDENTS BY TYPE AND CLASS

SOURCE OF VARIANCE	SUM OF SQUARES	DF	MEAN SQUARE	F	P
Main Effects	2.60	4	0.65	2.37	0.06
Type	0.66	2	0.33	1.20	0.31
Class	1.76	2	0.88	3.21	0.05
2-Way Interactions					
Type x Class	0.49	3	0.16	0.59	0.62
Explained	3.09	7	0.44	1.61	0.15
Residual	14.81	54	0.27		
Total	17.90	61	0.29		

CHAPTER V

SUMMARY AND CONCLUSIONS

The previous chapter contained the results of the study outlined in Chapter III. In this chapter, a summary of the study is presented, together with the major conclusions and implications to be drawn from it.

Summary

The purpose of this study was to examine the growth of critical thinking of senior high school students as a function of the number and kind of science courses that they completed in grades eleven and twelve.

At the conclusion of grade ten, a sample of two hundred twenty students was tested with the Watson-Glaser Critical Thinking Appraisal, Form YM. At the end of grade twelve, the sample number had been reduced, for reasons beyond the control of the experimenter, to a sample of sixty two students. These students were tested using the Watson-Glaser Critical Thinking Appraisal, Form ZM.

At the end of the data collection phase, the gain in critical thinking was expressed in terms of differences between pretest and posttest z-scores. The data were analysed by correlation and analysis of variance techniques. In this analysis, the subjects were classified on the bases of age, sex, class (science major, science minor and non-science) and type (university entrance, mixed and general). The effects of specific subject areas such as physics, chemistry, biology and mathematics were examined.

The results of the study point to several conclusions which merit discussion and from which implications for future educational practice may be drawn.

Conclusions

In this section, the results of the study are interpreted in terms of the the questions posed in Chapter I, pages four and five.

Question 1: Is there an increase in the critical thinking ability of senior high school students between grades ten and twelve?

The null hypothesis of no difference in mean gain in critical thinking of students between grade ten and twelve was rejected. Indeed, the mean gain observed in this study exceeded the normal two-year gain implied by Watson and Glaser. This finding, therefore, supports the data reported by Watson and Glaser (1964, Tables 5b and 5c).

Question 2: Is there a difference in the mean gain in critical thinking of students classified according to age and sex?

The null hypothesis of no difference in mean gain in critical thinking of students classified by age and sex could not be rejected. This fact is interpreted to mean that sex is not a significant variable in the growth in critical thinking. The study then supports that finding by Watson and Glaser (1964) but refutes that by Clark and De Roche (Belanger, 1969). The fact that this study dealt with mature students in comparable high school programs probably accounted for the finding that sex was not a significant variable.

The finding that age is not a significant variable supports the finding of Passmore (1969). On the other hand, the finding that there was a significant gain in z-scores between grades ten and twelve indicates that grade is a significant variable. In view of the fact that age and grade are usually found to be highly correlated variables (or at least confounded variables), this finding is unexpected. On the basis of these results, it can only be concluded tentatively that the increase in critical thinking between grades ten and twelve should be attributed to the effect of school. Thus, the question of what in the schooling process contributes to the growth of critical thinking really becomes crucial. This question is the focus of the remainder of the thesis.

Question 3: Do students high in critical thinking ability tend to select a greater number of specific science courses?

The null hypothesis of zero correlation between the Watson-Glaser pretest, z-scores and the subsequent number of science courses taken, whether university entrance, general or combined, was rejected in every case. For university entrance courses, the null hypothesis of zero correlation was rejected in favor of the alternative of a highly significant positive correlation. For general science courses, the null hypothesis of zero correlation was rejected in favor of a highly significant negative correlation. When no distinction was made between the type of course, the null hypothesis of zero correlation was rejected in favor of the alternative of a highly significant positive correlation. Here, the magnitude of the Pearson r was found to lie between that of the university entrance type and the general course type distinctions. Consequently, a tentative conclusion to be drawn from the results of the study is that those students endowed with more critical thinking ability do tend to choose more university entrance science courses in grades eleven and twelve than do those who do not. Alternatively, it is suggested that those students who are less endowed with critical thinking ability tend to enroll in the general science courses.

Question 4: Is there a difference in the mean gain in critical thinking of science and non-science students?

The null hypothesis of no difference in the mean gain in critical thinking of students classified on the basis of class (science major, science minor or non-science) was rejected at the .05 level of significance. An examination of the magnitude of the means of the treatment groups (See Table C-3, Appendix C) revealed that the science majors made the greatest mean gains in critical thinking. In contrast, the science minors made an even smaller mean gain than the non-science students. Since the number of science courses completed was the defining attribute of the levels of the variable class, this finding was unexpected. From these results, it may be concluded that critical thinking was developed to some degree by courses other than science. This hypothesis, however, did not come within the scope of the study and so was not tested.

Question 5: Among the science students, is there a difference in the mean gain in critical thinking of university entrance and general course students?

The finding of a significant positive correlation between the Watson-Glaser z-score gain and the number of university entrance science courses taken, as opposed to the finding of a significant negative correlation between the Watson-Glaser z-score gain and the number of general science courses taken, suggested that students taking university entrance courses not only start higher in critical thinking (See Table 4, page 32) but gain more than do those students taking general science courses. This finding was supported by Atwood (1967) who found that students who showed a preference for critical questioning of information gained more in critical thinking that those who showed a preference for memory of specific facts.

The null hypothesis of no difference in mean gain in critical thinking of students classified by type (university entrance or general science course) was tested, therefore, in the individual subject areas of physics, biology and chemistry. The null hypothesis was rejected for chemistry only. This finding is in contrast to those of George (1965), who found that students using the Molecules to Man version of BSCS biology scored significantly higher in critical thinking than those using the more traditional biology course materials, and of Henkel (1967), who found that the PSSC physics students had the more significant gains in critical thinking.

Taken together, these findings suggest that the nature of the subject content may be an important factor or alternatively, that specific types of programs may be interacting differentially with personality factors. Consequently, the null hypothesis of no difference in mean gain between students classified by type was tested in the area of mathematics. Here, the null hypothesis was rejected well beyond the .001 level of significance.

The association between the number of university entrance science courses and gain is both positive and cumulative. In the quest for increased critical thinking, this association has important implications for those planning the high school programs of individual students.

Question 6: Among university entrance and general course students, is there a difference in the mean gain in critical thinking of students who take and complete more science courses?

The null hypothesis of no difference in mean gain in critical thinking by students classified by type and class was rejected for the main effects of the variable class. An examination of the mean gains in critical thinking (Table C-8, Appendix C) shows a distinct difference in favor of the science majors. It is apparent that the findings of this study support the conclusion that the number of science courses is an important factor in the growth of critical thinking in students in the senior grades in the high school program.

Needed Research

It is clear from the evidence presented in this thesis that increases in grade level are parallelled by increases in critical thinking ability. Furthermore, within the limitations of the small number of subjects involved in the study, the importance of science courses as a factor in the growth of critical thinking in grades eleven and twelve is established. If one holds that the analysis of the data is appropriate and if one assumes that critical thinking is an important educational goal, then the subsequent replication of this study with a larger sample is desired. In addition, the scope of the investigation should be expanded to include all the major subject areas, or alternatively, to provide more specifically for the intervention effects of particular course participation.

If, in replication, the results of this study are supported, then the logical question to ask is "Why do science students have greater gains in critical thinking ability than non-science students?" One might suggest that the science courses content itself requires more complex thinking than the non-science course and as a result does enhance critical thinking. One could also suggest that students who enroll in the science courses are genetically more capable of critical thinking than non-science students. It is also possible that the gain in critical thinking of science students is a reflection of the attitudes toward education in general and science in particular. If this is so then a study incorporating the measurement of attitudes toward science could provide useful information.

It has been shown that the critical thinking ability of the teacher affects the critical thinking ability of the students (Chapter II). Do science teachers with greater academic science training necessarily induce the greater gains in critical thinking in their students in a fixed period of time?

Why is there a positive correlation between the gain in critical thinking and the number of university entrance science courses completed on the one hand and a negative correlation with the number of general science courses completed on the other? Might it not be possible that the degree and type of science experience interacted differently with some unidentified personality variables (cf. Hunt and Sullivan's Person X Environment X Behavior hypothesis)?

It has been suggested by Brown and Brown (1971) that establishing a favorable environment is a prime factor in the development of critical thinking ability. Do teachers establish an equally favorable atmosphere for university entrance and general science students? Or do they, assuming a lesser need or ability for critical thinking by these students, fail to provide the proper incentive? Can teachers be trained to act as a catalyst in bringing about such an atmosphere? Both Ausubel and Skinner have had something to say about the way in which such an atmosphere can be established.

It has been suggested that teaching by inquiry provides an environment in which the interpersonal relationships of the teacher and the learner enable students to obtain practice in critical thinking. In this environment, teachers are encouraged to relinquish a great deal of their traditional role of authority and to set up a learning environment in which the pupil is guided along through the processes of scientific problem solving or through question discussion sessions structured along the discovery pathway. Does immediate feedback in such situations, either by the teacher or other students, encourage more rigorous thinking and tend to lead pupils along the analytic pathway to better critical thinking?

Finally, if critical thinking is to be attained, evaluation must be in tune with this objective. Would the use of tests in which students are required to identify problems, to select and interpret pertinent information, to recognize assumptions, to formulate possible hypotheses and to draw valid conclusions and inferences, help foster critical thinking?

BIBLIOGRAPHY

BIBLIOGRAPHY

Atwood, Ronald K. Chem study achievement among groups classified by cognitive preference scores. Journal of Research in Science Teaching, 1967, 5, 154-159.

Ausubel, David P. An evaluation of the "Conceptual Schemes" approach to science curriculum development. Journal of Research in Science Teaching, 1965, 3, 254-264.

Ausubel, David P. An evaluation of the BSCS approach to high school biology. American Biology Teacher, 1966, 28, 176-186.

Belanger, Maurice. Learning studies in science education. Review of Educational Research, 1969, Vol. 39: No. 4, 377-395.

Burmester, Mary. Behavior involved in the critical aspects of scientific thinking. Science Education, December 1952, 259-263.

Burton, William; Kimball, Roland; & Wing, Richard. Education for effective thinking. New York: Appleton-Century Crafts, 1960, 29-451.

Bruner, Jerome S. The process of education. Cambridge, Mass.: Harvard University Press, 1960.

Brown, Stanley B. & Brown, Barbara. Suggested critical thinking and inquiry techniques in science for middle school teachers. School Science and Mathematics, 1971, 8, 731-736.

Charen, George. Do laboratory methods stimulate critical thinking? Science Education, 1970, Vol. 54: No. 3, 267-271.

D'angelo, Edward. Philosophers as critical thinking consultants. School and Society, 1966, 98.

Dressel, Paul. Critical thinking, the goal of education. National Education Association Journal, 1955, 41, 418-420.

Dressel, Paul & Mayhew, Lewis. General education explorations in evaluation. Washington, D.C.: American Council on Education, 1954, 174-207.

Dyck, Jacob. Changing enrolment patterns in science education. M.Ed. Thesis, University of Manitoba, October 1974.

Ennis, Robert H. A concept of critical thinking. Harvard Educational Review, 1962, 32, 81-111.

Gagné, Robert M. The learning requirements for inquiry. Journal of Research in Science Teaching, 1963, 1, 144-153.

George, Kenneth D. The effect of BSCS and conventional biology on critical thinking. Journal of Research in Science Teaching, 1965, 3, 293-299.

George, Kenneth D. & Dietz, Sister M. The relationship of teacher-pupil critical thinking ability. Science Education, 1968, Vol. 52: No. 5, 426-431.

Glaser, Edward M. An experiment in the development of critical thinking. Ann Arbor, Michigan: University Microfilms 1963, (Reprint of Contributions to Education No. 843), Teachers College, Columbia University, 1941.

Glass, Bentley. Theory into action - a critique. Science Teacher, May 1965, 32, 29-30.

Guilford, J. P. Three faces of the intellect. American Psychologist, 1959, 14, 469-479.

Guilford, J. P. & Hoepfner, L. Rep. psychology lab no. 30. Los Angeles: University of Southern California, 1963.

Haas, Sister Geraldine. A study of critical thinking and reading ability. Unpublished Doctoral Dissertation, Lordham University, 1963.

Henkel, E. Thomas. The effect of instruction in physics upon the critical thinking ability of undergraduate students. Journal of Research in Science Teaching, 1967, 5, 89-94.

Herron, J. Dudly. Evaluation and the new curricula. Journal of Research in Science Teaching, 1966, 4, 159-170.

Houle, C. O. Evaluation in the eight year study. Curriculum Journal, 1943, 14, 18-21.

Hunt, David E. & Sullivan, Edmund V. Between psychology and education. Hinsdale, Ill.: Dryden Press, 1974.

Inhelder, Barbel. Some Aspects of Piaget's genetic approach to cognition. In J. Cohen (Ed) Readings in psychology. London: Allan and Urwin, 1964, 85-103.

Johnson, D. M. Problem solving processes. Paper read at American Psychological Association. St. Louis, September 1963.

Kastrinos, William. The relationship of two methods of teaching to the development of critical thinking by high school students in advanced biology. Science Education, 1964, 48, 187-195.

Madison, John P. Critical thinking in the English classroom. English Journal, 1964, Vol. 60 (8), 1133-1140.

Manitoba High School Program of Studies Grades 9-12, 1975-76.

Morse, H. T. & McCune, G. H. _Selected items for the testing of study skills and critical thinking_ (third edition). Washington, D.C.: National Council for the Social Studies, 1957.

Passmore, John. On teaching to be critical. _The Concept of Education_, London: Routledge and Kegan Paul, 1967.

Rickert, Russel K. Developing critical thinking. _Science Education_, 1967, _51_, 24-27.

Schwab, Joseph J. _The teaching of science as enquiry_. (Edited by J. Schwab & Paul L. Brandwein). Cambridge, Mass.: Harvard University Press, 1962, 3-103.

Skinner, Ballow. The myth of teaching for critical thinking. _The Clearing House_, 1971, _45_ (6), 372-376.

Tate, Merle W. & Stanier, Barbara. Errors in judgment of good and poor problem solvers. _Journal of Experimental Education_, Summer 1964, _32_, 371-376.

Troost, Cornelius. Teaching critical thinking: Practical suggestions. _School Science and Mathematics_, 1971, 71 (7), 619-622.

Voss, Burton E. & Brown, Stanley B., eds. _Biology as an enquiry_. Saint Louis: C. N. Mosby Co., 1968.

Watson, Goodwin B. _The measurement of fairmindedness contributed to education_. New York: Bureau of Publication, Teachers College, Columbia University, No. 196, 1925.

Watson, G. & Glaser, E. M. _Watson-Glaser critical thinking appraisal manual_. New York: Harcourt, Brace, & World, 1964.

Watson, G. & Glaser, E. M. _Critical thinking appraisal form YM_. New York: Harcourt, Brace, & World, 1964.

Wertheimer, M. _Productive thinking_. New York: Harper & Row, 1945.

APPENDIX A

TABLE A-1

SUMMARY OF AGE AND MENTAL ABILITY CHARACTERISTICS OF THE CRITICAL THINKING APPRAISAL HIGH SCHOOL NORMATIVE SAMPLES

(cf. Table 5a Watson and Glaser, 1964)

	Grade 9		Grade 10		Grade 11		Grade 12		Grades 9-12	
	YM	ZM	YM	ZM	YM	ZM	YM	ZM	YM	ZM
Number in Sample	3037	3033	2947	2995	2406	2316	1800	1778	10,114	10,036
OTIS RAW SCORE										
Range	6-76	3-77	6-77	8-77	12-79	9-78	9-78	13-79	6-79	3-79
Mean	43.5	44.0	49.0	48.1	52.4	52.1	53.7	54.3	49.0	48.9
Standard Dev.	12.8	12.9	12.9	12.9	12.9	12.9	12.6	12.8	13.8	13.4
Correlation	.75	.70	.75	.73	.76	.73	.75	.75	.75	.73
OTIS IQ										
Range	67-144	65-142	64-139	66-138	70-139	67-138	67-136	71-138	64-144	65-142
Mean	107.1	107.5	110.0	109.1	111.8	111.4	112.0	112.6	109.8	109.7
Standard Dev.	13.4	13.5	13.4	13.2	13.0	13.0	12.7	12.8	13.4	13.4
CHRONOLOGICAL AGE*										
Range	11-2 to 21-6		12-7 to 23-1		11-8 to 22-7		11-8 to 22-1		11-2 to 23-1	
Mean	14-10		15-10		16-9		17-9		16-1	

* Age is given in years and months for total grade, both forms combined.

TABLE A-2

SUMMARY OF MENTAL ABILITY CHARACTERISTICS
OF THE CRITICAL THINKING APPRAISAL
COLLEGE NORMATIVE SAMPLES
(cf. Table 5d Watson and Glaser, 1964)

	LIBERAL ARTS FRESHMEN	COLLEGE SENIOR WOMEN
Number in Sample	5297	554
OTIS RAW SCORE:		
Range	15-80	34-79
Mean	58.7	60.2
Standard Dev.	9.8	9.7
Correlation with Form YM	.60	.66
OTIS IQ:		
Range	73-138	92-137
Mean	116.7	118.2
Standard Dev.	9.9	9.7
Correlation with Form YM	.60	.66

TABLE A-3

SUMMARY OF CRITICAL THINKING APPRAISAL, FORM YM
TOTAL RAW SCORES FOR THE NORMATIVE SAMPLES

Normative Sample	FORM YM TOTAL RAW SCORE					
	N	Mean	Range	S.D.	Rel.[a]	S.E.m.[b]
Grade 9	3,037	57.7	19.91	11.0	.85	4.2
Grade 10	2,947	61.7	17-93	11.0	.86	4.0
Grade 11	2,406	64.4	21-94	11.0	.86	4.0
Grade 12	1,800	65.6	23-96	10.9	.87	3.9
Grades 9-12	10,114	61.8	17-96	11.4	.86	4.3
Lib. Arts Fresh.	5,297	70.2	19-95	9.8	.85	3.8
College Seniors[c]	200	74.4	37-97	9.6	.85	3.7

[a] Odd-even split-half reliability coefficients corrected by Spearman-Brown formula.

[b] Standard errors of measurement computed from corrected split-half reliabilty coefficients.

[c] Random sample of 200 cases from a population of 554 senior women in ten liberal arts colleges (22)

APPENDIX B

INSTRUCTIONS FOR ADMINISTERING
THE WATSON GLASER CRITICAL THINKING APPRAISAL

A test booklet, an answer sheet, two soft-lead pencils and a good eraser are required for each person taking the test. The following directions are intended specifically for administering the test with IBM 805 answer sheets. Supplementary directions will be packaged with IBM 1230 or other types of answer sheets as they are published.

The examiner should plan to have at least 55 minutes available to take care of the actual working time and the time spent in giving directions, passing out materials, and other preliminary activity.

After distributing all the necessary materials, say:

"May I have your attention, please? Each of you has been given a test booklet, (a special pencil), and a separate answer sheet. Do not open the booklet or make any marks on the answer sheet until I tell you to do so. Now fill in your name and the other information called for on the left-hand side of the answer sheet. The date of testing is.... The form of the test you will be taking is Form... so circle the... on your answer sheet."

When all the information has been filled in on the answer sheet, say:

"Each test is preceded by its own directions. When I tell you to begin, read carefully the directions for the first test and study the sample questions until you know what you are to do. If you don't understand the directions, raise your hand and I will explain them to you. Do not ask questions about a test after you have started work on it. Don't make any marks on the test booklet.

"For each question, decide what you think is the best answer. Then record your choice by making a black mark in the appropriate space on the answer sheet. Always be sure that the answer space is numbered the same as the question in the booklet. Do not make any other marks on the answer sheet. If you change your mind about an answer, be sure to erase the first mark completely. You may answer a question even when you are not perfectly sure that your answer is correct, but you should avoid wild guessing. Do not spend too much time on any one question. When you finish a page, go right on to the next one. If you finish all the tests before time is up, go back and check your answers. Work rapidly and accurately.

("These tests are to be scored electrically, so failure to follow these instructions can reduce your score. Be sure to use only the special pencil to mark your answers.)

"You will be allowed 13 minutes for the first test. This is ample time for most of you to answer every question without hurrying if you do not take too long on any one question. When you finish Test 1, go right on to Test 2 without waiting.

"So that you will have a guide in spacing your time, I am going to stop any one of you who have not finished each test in the usual time and start you on the next test. Those who run a bit short of time on some tests may have time left at the end. When you finish Test 5, the last test, you can go back and answer any questions that you skipped, or check your answers to the other questions. If you finish a test before time is called, go on to the next test.

"Remember, you are to start reading the directions for Test 1 when I tell you to start and continue working through the successive tests until I tell you to stop. If you wish to change an answer, erase completely. Make no marks on the test booklet. Are there any questions before we begin?

"All right now, open your booklet and begin."

In order to insure that even the slowest persons attempt most of the items in each subtest, the examiner should note the starting time, successively add the time suggested below for each test, and as each finishing time arrives, say to the group: "If you are still working on Test...., stop and go to Test.... You may go back and finish later if you need more time." Since this is a power test and not a speed test, persons requiring more time should be given the opportunity to finish at their own pace.

Test	Suggested Time
1. Inference	13 min.
2. Recognition of Assumptions	6
3. Deduction	11
4. Interpretation	12
5. Evaluation of Arguments	8
Total	50 min.

Allow the group to continue working until more than 95 percent (all but one or two in a group of 30) have finished. A few slower ones may need more than the recommended time. This additional time may be added to the time allowed for the last test.

APPENDIX C

TABLE C-1

Z-SCORE DISTRIBUTION STATISTICS*

STATISTIC	FORM YM	FORM ZM	GAIN
MEAN:	-0.108	0.349	0.457
Mode	0.327	0.533	0.206
Median	-0.082	0.462	0.438
VARIANCE:	0.860	0.862	0.293
St. Dev.	0.928	0.928	0.542
ST. ERROR:	0.118	0.118	0.069
RANGE:	4.000	4.19	3.095
Minimum	-2.127	-1.94	-1.339
Maximum	1.873	2.248	1.756
KURTOSIS:	-0.729	-0.439	2.254
SKEWNESS:	-0.137	-0.216	-0.350
N	62	62	62

*In computing these z-scores, the means and standard deviations of the Watson-Glaser normative samples for grade ten (Form YM) and for grade twelve (Form ZM) were used. This procedure was adopted on the rationale that the study sample was representative of the Watson-Glaser sample (population). See Table 1, page 2.

TABLE C-2

WATSON-GLASER CRITICAL THINKING MEAN GAIN SCORES
BY AGE AND SEX

VARIABLE	N	UNADJ. DEV.	ETA	ADJ. DEV.	BETA
AGE:					
15	5	-0.05		-0.05	
16	46	0.06		0.06	
17	10	-0.32		-0.32	
18	1	0.51		0.56	
			0.29		0.29
SEX:					
Male	26	-0.05		-0.05	
Female	36	0.04		0.04	
			0.08		0.08

Multiple R Squared = 0.090
Multiple R = 0.300
Grand Mean = 0.46

TABLE C-3

WATSON-GLASER CRITICAL THINKING MEAN GAIN SCORES
BY AGE, SEX AND CLASS

VARIABLE	N	UNADJ. DEV.	ETA	ADJ. DEV.	BETA
AGE:					
15	5	-0.05		-0.20	
16	46	0.06		0.06	
17	10	-0.32		-0.22	
18	1	0.51		0.56	
			0.29		0.25
SEX:					
Male	26	-0.05		-0.08	
Female	36	0.04		0.06	
			0.08		0.12
CLASS:					
Non-Science	6	0.13		0.03	
Science Minor	34	-0.16		-0.15	
Science Major	22	0.21		0.22	
			0.33		0.32

Multiple R Squared = 0.180
Multiple R = 0.425
Grand Mean = 0.46

TABLE C-4

WATSON-GLASER CRITICAL THINKING MEAN GAIN SCORES
FOR PHYSICS STUDENTS BY SEX AND TYPE

VARIABLE	N	UNADJ. DEV.	ETA	ADJ. DEV.	BETA
SEX:					
Male	12	-0.13		-0.15	
Female	12	0.13		0.15	
			0.25		0.29
TYPE:					
Physical Science	10	-0.12		-0.15	
Physics	14	0.08		0.11	
			0.18		0.23

Multiple R Squared = 0.114
Multiple R = 0.338
Grand Mean = 0.56

TABLE C-5

WATSON-GLASER CRITICAL THINKING MEAN GAIN SCORES
FOR BIOLOGY STUDENTS BY SEX AND TYPE

VARIABLE	N	UNADJ. DEV.	ETA	ADJ. DEV.	BETA
SEX:					
Male	17	-0.14		-0.10	
Female	24	0.10		0.07	
			0.19		0.15
TYPE:					
General Biology	16	-0.14		-0.11	
U.E. Biology	25	0.09		0.07	
			0.19		0.15

Multiple R Squared = 0.057
Multiple R = 0.238
Grand Mean = 0.41

TABLE C-6

WATSON-GLASER CRITICAL THINKING MEAN GAIN SCORES
FOR CHEMISTRY STUDENTS BY SEX AND TYPE

VARIABLE	N	UNADJ. DEV.	ETA	ADJ. DEV.	BETA
SEX:					
Male	10	-0.15		-0.12	
Female	13	0.11		0.09	
			0.29		0.24
TYPE:					
Chemistry	8	-0.34		-0.32	
Phy. Science	15	0.18		0.17	
			0.54		0.52

Multiple R Squared = 0.352
Multiple R = 0.593
Grand Mean = 0.69

TABLE C-7

WATSON-GLASER CRITICAL THINKING MEAN GAIN SCORES
FOR MATHEMATICS STUDENTS BY SEX AND TYPE

VARIABLE	N	UNADJ. DEV.	ETA	ADJ. DEV.	BETA
SEX:					
Male	23	-0.10		-0.11	
Female	29	0.08		0.09	
			0.16		0.18
TYPE:					
General Math	35	-0.19		-0.19	
U.E. Math	17	0.38		0.39	
			0.49		0.49

Multiple R Squared = 0.266
Multiple R = 0.516
Grand Mean = 0.44

TABLE C-8

WATSON-GLASER CRITICAL THINKING MEAN GAIN SCORES FOR STUDENTS BY TYPE AND CLASS

VARIABLE	N	UNADJ. DEV.	ETA	ADJ. DEV.	BETA
TYPE:					
U.E.	48	0.06		0.05	
Mixed	2	-0.44		-0.46	
General	12	-0.16		-0.11	
			0.22		0.19
CLASS:					
Non-science	6	0.13		0.16	
Science minor	34	-0.16		-0.15	
Science major	22	0.21		0.19	
			0.33		0.32

Multiple R Squared = 0.145
Multiple R = 0.381
Grand Mean = 0.46