AN EXAMINATION OF THE CRITICAL THINKING ABILITY OF ENTERING FIRST YEAR UNIVERSITY STUDENTS AND PROSPECTIVE SOCIAL STUDIES TEACHERS NEARING THE COMPLETION OF THEIR HIGH SCHOOL TEACHER PREPARATION PROGRAMME AT MEMORIAL UNIVERSITY OF NEWFOUNDLAND

HARVEY STANFIELD RICE

Memorial University of Newfoundland

An Examination of the Critical Thinking Ability of Entering First Year University Students and Prospective Social Studies Teachers Nearing the Completion of Their High School Teacher Preparation Programme at Memorial University of Newfoundland

by

© Harvey Stanfield Rice, B.A., B.Ed.

A Thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Education

Department of Curriculum and Instruction

November, 1986

St. John's Newfoundland

ISBN 0-315-37014-9

ABSTRACT

The purpose of the present study was to determine the critical thinking abilities of entering first year university students and that of prospective social studies teachers nearing the completion of their high school teacher preparation programme at Memorial University of Newfoundland. Several variables, including university training, were explored to determine their relationship to critical thinking ability.

Since social studies is often taught by graduates from disciplines not included in the social studies, it was necessary to assess the critical thinking ability of students enrolled in a variety of specializations. One hundred and sixty two students enrolled in appropriate methods courses, representing four areas of concentration, were tested.

The sample of first year students (N = 164) was randomly selected, by class, from those enrolled in English 1000. This course was selected since it is a requirement for all first year students.

The Ennis-Weir Critical Thinking Essay Test was used as an instrument to evaluate critical thinking ability. This test purports to measure a "person's ability to appraise an argument" rather than the conclusion emanating from the argument.

Analysis of Variance was used to evaluate results. An examination of scores obtained by first year students

indicated an interaction between sex and urbanness significant at the .07 level and an interaction between sex and maturity significant at the .09 level. Analysis of Variance for simple main effects revealed that the score of rural males was significantly better ($p < .03$) than the score of rural females. In addition, regular males (those not classified as mature students) achieved significantly higher ($p < .04$) scores than did regular females.

ANOVA results of senior students indicated that no significant differences were detected among students enrolled in the high school teacher preparation programme according to sex, year, or area of specialization. However, significant interaction at the .09 level was detected between year and the number of philosophy courses.

When comparisons were made between the total scores obtained by first year students and those achieved by fourth and fifth year students, significant differences were detected. The differences exist primarily between first year students and those senior students specializing in English, social studies, or science.

A significant outcome of the present study was an extensive evaluation of the Ennis-Weir Critical Thinking Essay Test. Several questions were raised which could have serious implications for the extensive use of the instrument.

ACKNOWLEDGEMENTS

The writer wishes to express his sincere thanks and appreciation to his supervisors, Dr. Frank Cramm and Dr. Stephen P. Norris. Throughout the research and writing stages, the supervisors were always available to give the guidance, encouragement and constructive criticism so essential for the completion of a thesis.

Encouragement also came from my wife, Barbara. Her willingness to make sacrifices and her faith in my abilities will always be appreciated.

As always, my two daughters, Jennifer and Heather, provided inspiration.

Finally, thanks is also due to Dr. Eric Weir, co-designer of the Ennis-Weir Test, for his willingness to score samples that enabled inter-rater reliabilities to be established.

This Thesis is dedicated to the Memory of three of my Senior High School Students who died in a tragic accident on June 8, 1985:

Breene Saunders

Nancy Saunders

Jeff Stride

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

OVERVIEW

The purpose of the present study is to ascertain the critical thinking ability of entering first year university students; the critical thinking competency of beginning social studies teachers; and the influence of university training on the development of critical thinking ability. In addition, several variables which may influence the level of critical thinking ability will be explored. In order to survey beginning social studies teachers, the critical thinking ability of students enrolled in a variety of specializations will be examined. This is necessary since social studies is often taught by graduates from disciplines not included in the social studies.

This study is motivated by two beliefs. First, since the application of critical thinking skills should form an integral part of the daily activities of us all, it is important to establish the level of critical thinking ability of high school and university graduates and some factors which may contribute to such levels. Second, the investigation into the critical thinking competence of beginning high school social studies teachers is of particular significance because of the emphasis in that set of subjects on teaching critical thinking. To help students acquire critical thinking skills, teachers must possess such skills themselves.

The evaluation instrument, The Ennis-Weir Critical Thinking Essay Test: An Instrument for Testing and Teaching (Ennis and Weir, 1985a), was used as the measure of critical thinking ability for this study. The instrument is based on Robert H. Ennis' conceptualization of critical thinking which involves "reasonable reflective thinking that is focused on deciding what to believe or do" (Ennis, 1985c). This concept of critical thinking was adopted for this study.

Background

The development of critical thinking skills has long been acknowledged as a primary objective of education. The intellectual roots of the critical thinking movement can be traced back to the early Greek philosophers. Aristotle (384-322 B.C.) perceived living as being essentially composed of three basic activities: wanting, thinking, and doing. Philosophers throughout history such as St. Augustine, St. Thomas Aquinas, René Descartes, Immanuel Kant, John Locke, and David Hume have stressed the importance of critical thinking to society.

Despite the fact that critical thinking has nearly always been stated as a primary objective of education, the goal of critical thinking has not received much priority in actual instruction. However, it appears that the critical thinking movement has been revitalized due to

impetus from two important sources. First, in 1980 the Rockefeller Commission on the Humanities recommended that the U.S. Office of Education include critical thinking in its definition of the basic skills. Second, since 1983 the nineteen campus California State University requires all students to complete a course in critical thinking in order to graduate. This graduation requirement is intended to provide students with

> ... an understanding of the relationship of language to logic, leading to the ability to analyze, criticize, and advocate ideas, to reason inductively and deductively, and to reach factual or judgmental conclusions based on sound inferences drawn from unambiguous statements of knowledge or belief. (cited in Paul, 1984, p. 5)

In the United States many universities, community colleges and high schools reacted to this development by instituting programs of their own in critical thinking. While the movement has not enjoyed the same momentum in Canada, there are indications of increased awareness, interest and study throughout the country. Two international conferences on critical thinking have been sponsored by the University of Windsor under the leadership of Anthony Blair and Ralph Johnson. These philosophers have also been responsible for the publication of a new journal Informal Logic, which is

devoted exclusively to critical thinking issues. The developmental work of Norris and King (1983) was also a significant contribution to the field. The prominence of critical thinking among Canadian educators was also highlighted by a recent edition of the History and Social Science Teacher (March, 1986) devoted to this topic. In addition, the University of Windsor offers a Master's degree programme in critical thinking. Only recently, (September, 1986), Memorial University gave senate approval for an optional course in critical thinking for students in the Master of Education degree program. Although the current emphasis on critical thinking appears to be primarily an academic issue among educators and philosophers at the university level, some critical thinking programs such as de Bono's Cognitive Research Trust (CoRT) appears to be gaining acceptance in some areas at the school level.

Statement of the Problem

Results of studies into thinking abilities of students have led many educators (Aylesworth and Reagan, 1969; Beyer, 1985b; Hodgetts, 1968; and Norris, 1985b) to state or infer that students are still taught what to think rather than how to think. There are many factors which account for the discrepancy between the stated objectives and the emphasis of instruction. Wright (1977), and Beyer (1985a) contend that instruction in critical thinking does not take place because the social

studies specialists do not have clear conceptions in their own minds as to the exact meaning of critical thinking. Anderson (1942) recognized this in the following:

> Social studies teachers have long accepted critical thinking as an important and desirable outcome of instruction. Or perhaps it would be more accurate to say that ... they have accepted critical thinking in principle without bothering to define the term precisely or to do much by way of direct instruction to see that this goal is achieved. (p. v)

Parsons and Shaftel (1967) concluded from a study on teaching behaviors that "though the ... teachers were able to articulate the professional ideology regarding thinking, they had only the vaguest notion of what thinking is" (p. 127). When compared to Anderson's (1942) assessment the statement reveals that little had changed in 25 years.

Gray (1969), Henderson (1972), Woods and Walton (1974), Beyer (1984b, 1985a) and Unks (1985) also suggest that the lack of teacher knowledge of critical thinking is a major factor resulting in little instruction in this area.

Other factors include a lack of appropriate curriculum materials (Beyer, 1984a, 1985a; Wright and LaBar, 1986; Crocker and Riggs, 1979) and a curriculum

suffering from "skills overload" (Beyer, 1984a). With students being bombarded with literally dozens of skills at each grade level, the teacher finds it difficult to determine the priority of critical thinking skills. These factors, and others, militate against the teaching of this important skill.

Within this background, the detailed purposes of this study are conceived as follows:

(1) To determine the critical thinking competence of beginning high school social studies teachers. To do this, it was necessary to evaluate the critical thinking ability of students enrolled in a variety of specializations since social studies is often taught by graduates from disciplines not included in the social studies. Students nearing the end of their high school teacher preparation programme and enrolled in appropriate instructional methods courses were tested. Several factors including sex, year, area of study and number of philosophy courses, which might influence critical thinking ability were examined.

(2) To determine the critical thinking ability of first year students attending Memorial University. Several variables including sex, maturity, and urbanness were explored to determine their influence on critical thinking ability. Results from this group may give some indication of the success of the high school programme, which includes a core of at least four social studies courses, in the promotion of critical thinking skills.

(3) To make comparisons between first year and senior students to determine if university training is related to critical thinking ability.

Rationale of the Study

The philosophical basis of education in Newfoundland, namely the Aims of Education in Newfoundland and Labrador (1984), recognizes the importance of the critical thinking component to the educational process. The document states that the individuals who have achieved their fullest and best development are those who, among other things, "have minds whose critical and other faculties are so developed and trained as to enable them to cope successfully with the varied problems and situations that they may be expected to encounter" (p. 3). Such a philosophy suggests that one of the fundamental objectives of education is to provide opportunities for the development of the pupil's abilities to think critically.

In 1967, the Royal Commission on Education and Youth was established to investigate the status of education in Newfoundland. While many radical changes were advocated through its 340 recommendations, the Commission reiterated the Province's Aims of Education. The report not only reflected the opinions of its twelve authors, but also the views of the Newfoundland citizenry. According to the Commission, the onus is on the school to "... produce

responsible, well-informed thinking citizens with the intellectual, social, emotional, and moral qualities necessary for successful living, and for the successful growth of society" (vol. 1, p. 144).

In the view of the Commission the traditional approaches to instruction, which often stress the mere memorization of facts, should no longer be tolerated. The Commission suggested that instruction be aimed at understanding, critical analysis and judgment. Only then would students develop thinking, reasoning, and creative abilities.

In a study conducted by Warren (1978) concerning public attitudes toward education in Newfoundland, he asked respondents to indicate both the quality they consider most important in the overall development of the child and the quality most neglected by the school today. The answer to both questions was "learning to think for oneself".

A second survey of public opinion (Warren, 1983) indicated that the Newfoundland public felt that teaching students to think was the most important function of elementary schools and 58% of respondents stated that high schools should place more emphasis on this important goal.

Given that the school is the only organization which directly and systematically concerns itself with the intellectual development of the individual, Crocker and Riggs (1979) suggest that the first aim of education

should be to allow students to reach their highest level of intellectual achievement. They sub-divide the domain of intellectual development into four components: skills, knowledge, analysis and intellectual independence. Regarding intellectual independence they say that "each child should be capable of making judgments on the basis of information given and of recognizing the consequences and limitations of such judgments, particularly when they are based on incomplete evidence" (p. 28). Intellectual independence, according to Crocker and Riggs, is simply a restatement of "learning to think for oneself". In fact, both are important aspects of critical thinking as defined in this study.

While critical thinking is regarded as an overall general educational objective, social studies educators perceive it as one of their prime goals. Fenton (1967) states that social studies "should help each student develop to the limit of his ability into an independent thinker and a responsible citizen of a democratic society". Other writers such as Newman and Oliver (1970); Paul (1984), and Unks (1985) acknowledge the strong relationship between effective citizenship and the ability to think critically.

A study by Guyton (1984) re-affirms this position. Her study was designed to measure the relationship between critical thinking and political participation. A four stage conceptual model was developed to ascertain the

relationship between personality variables (self-esteem, personal control, and political efficacy), critical thinking, democratic attitudes and political participation. Various instruments were used to measure these variables. The model, used with 118 undergraduate and graduate students, reveals that critical thinking positively affects personal control, political efficacy, and democratic attitude. In turn, political efficacy and democratic attitudes were found to positively affect political participation. The development of critical thinking skills thus appear to contribute to the development of responsible and participating citizens.

The social studies educators in Newfoundland appear to recognize that they have a major responsibility in contributing to the development of youth into rational and responsible citizens. The Master Guide for Social Studies K-XII of Newfoundland and Labrador has one overall goal: the development of person-centered and citizen-centered adults. To achieve this goal, the recommended underlying methodology of social studies is to focus on the development of critical thinking skills.

Many educational theorists and philosophers, outside the Newfoundland context, have endeavoured to present a rationale for the inclusion of critical thinking skills as one of the prominent goals of the educational process. For example, Scheffler (1973) maintains that "critical

thought is of the first importance in the conception and organization of educational activities".

Siegel (1980), a student of Scheffler's, has expounded three major reasons as to why critical thinking should be considered as a worthwhile educational ideal. First, critical thinking is relevant to and has implications for the ethics of education. The ethics consideration arises with respect to the manner of teaching and the learner's moral education. Educators must ensure that instructional methods meet certain moral standards and endeavour to contribute to the moral education of the learner. This is a tremendous responsibility although how it is to be accomplished is somewhat unclear. However it is done, according to Siegel "we must, if we are to conduct our interpersonal affairs morally, recognize and respect the fact that we are dealing with other persons who as such deserve respect" (p. 13). This respect for persons implies that it is only morally right for teachers to acknowledge the student's right to question, to challenge, and to demand reasons. If students become accustomed to questioning, challenging, and seeking reasons, the basis is laid on which they can reasonably decide what to do or believe.

Second, critical thinking is an educational ideal because it is essential in order to prepare for one's adult life. Toffler (1970) states that "the technology of tomorrow requires ... [people] who can make critical

judgments, who can weave their way through novel environments, who are quick to spot new relationships in the rapidly changing reality" (p. 402).

Scheffler (1973) asserts the view that to train students to become critical thinkers is to "encourage them to ask questions, to look for evidence, to seek and scrutinize alternatives, to be critical of their own ideas as well as those of others" (p. 143). Hitchcock (1983) indicates that practice in such skills protects people from being seduced by rhetoric, propaganda, or advertising. In addition, such skills enable people to make wise decisions and participate constructively in the democratic process. In essence, this is the rationale as envisioned by the authors of the Master Guide for Social Studies, K-XII in Newfoundland and Labrador.

Third, critical thinking is an educational ideal because it is an embodiment of rationality (Siegel, 1980; McPeck, 1981). Rationality is viewed as thought in which reason predominates. A student must be made aware that justifications for various decisions are needed, because in order to become a rational thinker the student must have reasons to support conclusions. A rational thinker must also think according to rules or principles which are justifiable. Only then can one rationally conceptualize the relationship among the various reasons and evaluate decisions on their own merit.

According to Scriven (1985) any system of education which does not regard training in critical thinking abilities is guilty of "culpable negligence or worse", because the survival of our democratic society depends upon citizens who are competent critical thinkers. Schools must provide an environment conducive to the development of the necessary thinking skills.

Clearly then, educators must endeavour to develop within their students the necessary understandings, values and related social studies skills. Only then will they be able to realize their own potential and participate fully and constructively in society. Newman and Oliver (1970); Guyton (1984); Paul (1984) and Glaser (1985) maintain that this is not possible without the ability to think.

Apart from a study by Caravan (1979), little or no research in critical thinking abilities has been conducted with respect to the Newfoundland educational scene. The results of the present study may give some indication of the extent to which teachers entering the teaching profession are qualified to teach the thinking skills deemed important by the Aims of Education and the Master Guide for Social Studies K-XII of Newfoundland and Labrador. The results may have implications for teachers, the Department of Education, curriculum planners, as well as Memorial University. Depending on the results, the study may also serve as a starting point for further

investigation into critical thinking in other areas of the educational environment.

CHAPTER II

REVIEW OF LITERATURE

The expressed goal of critical thinking has not been translated into classroom activity despite its prominence in social studies literature and curriculum guides. Since 1980 there has been a profusion of articles and journals devoted exclusively to this topic. Paul (1985c) conducted an ERIC computer search and identified 1,894 articles written about critical thinking since 1978. An effort will be made in this chapter to focus mainly on selected writings of those with extensive knowledge or research experience in the critical thinking field. In this chapter several interpretations of the nature of critical thinking will be explored and a summary of critical thinking research in the social studies will be presented.

The Nature of Critical Thinking

We live in a technological society in which knowledge is expanding at a phenomenal rate. Since the primary purpose of schools is to educate students on how to function in such a society, it is incumbent upon them to provide learning experiences conducive to obtaining the necessary skills which would enable them to think for themselves. However, before a researcher can endeavour to evaluate the current level of critical thinking of

students an understanding of the nature of critical thinking is essential. Such an understanding could provide a logical basis on which a decision can be made involving which concept of critical thinking to adopt. Such a decision will influence the structure of the researcher's study.

A cursory review of the literature suggests that there is a great diversity of opinion on what exactly constitutes critical thinking and how it should be taught. As a result of this diversity teachers have been exposed, and often bombarded, with competing definitions of critical thinking. To some educators (Sanders, 1966), critical thinking includes all thought processes beyond the memory category of Bloom's Taxonomy of Educational Objectives: Cognitive Domain (1956). Others (Oliver and Shaver, 1968; Maxim, 1977) view only the evaluation level of the same taxonomy as being involved with the critical thinking process. Still others (Wilen, 1985) view critical thinking as involving the analysis, synthesis, and evaluation levels of Bloom's Taxonomy.

Other social studies educators such as Ponder and Davis (1982) view critical thinking and inquiry as being synonymous. Fraenkel (1980) associates critical thinking with decision making whereas Anderson (1942) equates it with problem solving. While many (Morse and McCune, 1940; Carpenter, 1963; Fair and Shaftel, 1967; Kurfman, 1977) agree with Anderson, others are diametrically opposed.

According to Allen and Rott (1969), critical thinking and problem solving are definitely not the same. They state:

> Critical thinking ... begins with a previous claim, conclusion or product and considers the question, "Of what truth or worth is it?". Problem solving, on the other hand, begins with a perceived problem and asks, "How might this difficulty be resolved?". (p. 2)

Black (1953) and Werkmeister (1957) view critical thinking as the application of principles of logic. In 1962, Ennis (1962) regarded critical thinking as the "correct assessing of statements" whereas de Bono (1984) associates it with "spotting of faults". McPeck's (1981) conception of critical thinking is "the appropriate use of reflective skepticism within the problem area under consideration" (p. 7).

Paul (1982) observes that critical thinking can be viewed in two senses: a weak sense and a strong sense. In the weak sense, a series of skills is used by the thinker to discover mistakes in reasoning. In the strong sense, comprehensive thinking skills are emphasized to develop a "free, rational and autonomous mind" (p. 5). In addition, strong sense critical thinkers are not only able to gain fundamental insights into an issue but are able to do so while being cognizant of their own egocentric and sociocentric viewpoints (Wright & LaBar, 1986).

Any effort at consensus is further complicated by the fact that educators and philosophers often change or modify their original positions. For example, Beyer (1977) states:

> Skill of distinguishing between statement of provable fact and statement of personal opinion constitutes one analytical skill which is indispensable for the intelligent use of information. (p. 38)

By 1985, however, Beyer (1985a) seems to suggest that this skill is not really fundamental because performing the operation is often difficult and unsuccessful due to ambiguity in meanings of terms (p. 274). In addition, the concept of critical thinking - envisioned by Ennis(1962) as "the correct assessing of statements" differs from his 1985 definition in that in 1985 he was including not just the assessment of statements but some judgments as to what one should believe or do, and the dispositions required to make such judgments.

After conducting a review of critical thinking literature, Feely (1976) concluded that the various interpretations of the nature of critical thinking could be categorized into two paradigms: the mental paradigm and the logical paradigm. The characteristics of each are summarized in Table 1.

Table 1

Mental and Logical Paradigms: A Comparison

Mental Paradigm	Logical Paradigm
1) Critical thinking is distinct from other thinking processes. It can be stimulated and manifestations observed.	1) Complex tasks or judgments can be reduced and analyzed into lists of sub-tasks. Critical thinking is only umbrella term under which a variety of activities are subsumed.
2) Critical thinking is not taught but rather stimulated. Stimulation results from asking questions requiring higher mental processes than factual recall. Questions stimulate students to think.	2) Students are not taught to think by inducing thinking but rather the grounds on which judgments can be made are taught.
3) Resolution of a problem is based upon how one feels.	3) Resolution of a problem is based upon a priori standard or criteria.
4) Score from measures of critical thinking ability such as Watson-Glaser Thinking Appraisal are recorded as a single score.	4) Scores on measures of critical thinking ability are broken down into component parts.

Adapted from: Feely, T. (1976) Critical Thinking: Toward a Definition, Paradigm and Research Agenda. Theory and Research in Social Education. Vol. IV, No. 1: 1-19.

The mental paradigm views critical thinking as a unified general mental process. Critical thinking is viewed as a type of thought or mental process which, although not observable, can be stimulated by using higher order questions as outlined in Bloom's Taxonomy of Educational Objectives: Cognitive Domain. Higher order questioning requires students to engage in thinking processes not utilized by using simple recall or knowledge questions. The Caravan thesis cited earlier is an example of a study based on the mental paradigm. Others will be examined in this chapter under the section entitled Critical Thinking in The Social Studies.

Researchers who view critical thinking in the logical paradigm perspective (Ennis, 1980; Paul, 1982; Norris & King, 1983) recognize that this perspective itself has two dimensions - a logical one requiring a diversity of skills and a dispositional one which emphasizes such dispositions as open-mindedness, considering alternatives, seeking reasons, and trying to be well-informed. Ennis (1985b) has listed five major categories of skills and some 13 dispositions which characterize good critical thinkers.

The famous Milgram experiment (Milgram, 1963) illustrated the need for such a two dimensional approach. The experiment demonstrated that there is more involved in critical thinking than having the ability to solve certain problems or being in possession of certain moral

principles. His research involved 40 subjects who were ordered to administer electrical shocks to a learner (actor) in another room whenever the learner failed to give a correct response to a question. The voltage readings ranged from 15 to 450 volts. Subjects were made aware of increased voltage by labels marked slight shock, moderate shock, strong shock, very strong shock, intense shock, extreme intense shock, severe shock and the letters XXX on the last two switches. The subjects were ordered to push the next highest switch each time the learner gave an incorrect response. Subjects were made aware, through a window, of the agony caused by the shocks. Despite this fact, 26 of the 40 subjects administered the maximum voltage to the learner (actor) on orders from the experimenter. Although all of the adult subjects were aware that such action was immoral, 65 percent were not willing to act in accordance with the dictates of their conscience. Norris (1985b) maintains that "no matter what level of critical thinking skill a person possesses, it is of no practical benefit unless the person is disposed to use these skills when they are appropriate" (p. 44).

The need for a second dimension becomes evident. Rational thinkers who possess the dispositions, sensitivities, and tendencies comprising the second dimension have often been referred to as having the critical spirit (Paul, 1982; Siegel, 1980; Norris, 1985b).

Many of the attributes of the critical spirit were outlined by D'Angelo (1971) who saw the following attitudes as prerequisite for the development of critical thinking:

1. Intellectual curiosity. Seeking answers to various kinds of questions and problems. Investigating the causes and explanation of events; asking why, how, who, what, when, and where.
2. Objectivity. Using objective factors in the process of making decisions. Relying on empirical evidence and valid arguments, and not being influenced by emotive and subjective factors in reaching conclusions.
3. Open-mindedness. A willingness to consider a wide variety of beliefs as possibly being true. Making judgments without bias or prejudice.
4. Flexibility. To be willing to change one's beliefs or methods of inquiry. Avoiding steadfastness of belief, dogmatic attitude, and rigidity. A realization that we do not know all the answers.
5. Intellectual skepticism. Postponing the acceptance of a conclusion as being true until adequate evidence is presented.

6. Intellectual honesty. The acceptance of statements as being true when there is sufficient evidence, even though it negates some of our cherished beliefs. To avoid slanting certain facts to support a particular position.
7. Being systematic. Following a line of reasoning consistently to a particular conclusion. Avoiding irrelevancies that stray from the issue being argued.
8. Persistence. To persist in seeking ways of resolving disputes. Supporting certain points of view without giving up the task of finding evidence and arguments.
9. Decisiveness. To reach certain conclusions when the evidence warrants it. To avoid unnecessarily drawn out arguments, snap judgments, and delaying reaching decisions until all the necessary information is obtained.
10. Respect for other viewpoints. A willingness to admit that you may be wrong, and that other ideas you do not accept may be correct. Listening carefully to another point of view and responding accurately to what has been said. (p. 7-8)

Ennis (1980) has modified D'Angelo's original list and responds that rational or critical thinkers are those who have the inclination to:

1. exercise the proficiency they possess;
2. take into account the total situation;
3. be well-informed;
4. demand as much precision as the subject matter permits;
5. deal with the parts of a complex situation in an orderly fashion;
6. consider seriously other points of view than one's own;
7. withhold judgment when the evidence and/or reasons are insufficient;
8. take a position (and change the position) when the evidence and reasons are sufficient to warrant so doing;
9. accept the necessity of exercising informed judgment; and
10. exercise good judgment. (p. 6)

If critical thinkers possess these dispositions they will have the tendency to act in accordance with their ability.

In addition to possessing certain abilities and having a critical spirit, a critical thinker must also have sound knowledge of the subject matter (Norris, 1985b; Ennis, 1985a; McPeck, 1981). Norris (1985b) states:

> A set of critical thinking skills, however well developed, cannot compensate for lack of knowledge in the area of question. The

> application of critical thinking principles involves a competence over and above knowledge of the principles themselves. (p. 44)

D'Angelo (1971), however, feels that attitudes and a knowledge and application of certain thinking skills are the only factors that occur in all areas of critical thinking. Knowledge of the subject area is not regarded as prerequisite for critical thinking. He observes:

> A knowledge of the subject area in which thinking occurs in often a necessary condition for the development of critical thinking. For example, certain knowledge is needed to determine whether a particular statement is a fact or an opinion. However, specific knowledge in a particular area is not always necessary in order to apply critical thinking skills. (p. 5)

An even more radical stance is adopted by de Bono (1985) who insists that training in thinking skills should not be dependent on the prior acquisition of specific knowledge. The de Bono critical thinking program (Cognitive Research Trust) is designed in such a way as to permit all students, regardless of intelligence, age, abilities and cultural background, to begin on an equal level (p. 366).

If, as Norris states, a sound knowledge base is required for productive thinking, the question arises as to how instruction should take place. McPeck (1981), Paul

(1982), Beyer (1984b), and Bereiter (1984) argue in favor of thinking skills being taught in traditional subject areas rather than as a separate subject. The latter position is advocated by de Bono (1985) and Glaser (1985). Norris (1985a) and Ennis (1985a) appear to remain uncommitted as to the most suitable means of instruction. Both acknowledge that there are general principles that appear to cross subject boundaries. Ennis (1985a) provides the following examples:

1. A person's having a conflict of interest is a ground for regarding that person's claim with greater suspicion than would otherwise be appropriate.
2. It is a mistake to misdescribe a person's position, and then attack the position as if it actually were the person's position (the "straw-person" fallacy).
3. Given an "if - then" statement denial of the consequent implies the denial of the antecedent.
4. The ability of a hypothesis to explain or help explain the facts lends support to the hypothesis, if the hypothesis is not otherwise disqualified. (p. 29)

As stated previously, great diversity of opinion exists regarding the nature of critical thinking. This diversity is not confined solely to the exact meaning of the term but extends to other areas as well - whether the

concept falls into a mental or logical paradigm, and whether or not specialized knowledge is prerequisite for the development of critical thinking skills.

It was acknowledged in Chapter One that Ennis' concept of critical thinking, representative of the logical paradigm, has been adopted for the present study. This decision was based on the belief that if students are to learn how to think for themselves the mental paradigm perspective of critical thinking is inadequate. It is not enough to simply expose students to higher level questions with the assumption that by doing so critical thinking will occur. Feely's (1976) evaluation of research and literature supports this position. He concluded that "the weight of both evidence and argument point toward the logical paradigm ... as clearly the most reasonable for both research and instruction in social education" (p. 11). This paradigm provides the criteria by which decisions and beliefs can be evaluated. Feely implies that the logical paradigm allows for a multi-stage approach to the teaching of critical thinking whereas the mental paradigm offers little guidance other than providing higher order questions. In addition, the structure of the logical paradigm allows for a concept-oriented curriculum. If the concepts are included in the curriculum and taught systematically the structure facilitates the acquisition of the ability by students to decide for themselves what to believe or do.

Critical Thinking in the Social Studies

The present study has tended to focus on social studies since the teaching of critical thinking is one of the fundamental aims in this area. As already indicated, there is a lack of consensus within the field as to the nature of critical thinking. The review of social studies literature in this section illustrates the two-paradigm approach to critical thinking within the discipline. The results of research studies examined will allow for a more realistic evaluation as to which perspective appears to be most suitable for the promotion of critical thinking skills in social studies. This evaluation will also provide further insight and justification for the adoption of the logical paradigm approach for the present study.

Social studies educators started to produce lists of skills covering critical thinking in their discipline as early as 1940 (Morse and McCune, 1940). Their National Council for the Social Studies publication identified 17 separate skills involved in the critical thinking process. This work provided the foundation for other social studies educators interested in the promotion of critical thinking. Educators such as Fraser and West (1961), Hudgins (1977), Beyer (1985a), and others have delineated lists of critical thinking skills, each reflecting to some extent the work of Morse and McCune.

Although the early emphasis by social studies educators was on skill development, much of the work

simply reflected lists of skills and descriptions of their importance. Rarely were the attributes of the various skills specified. However, the publication of Ennis' (1962) paper on critical thinking marked an important development in the field. According to Ennis, critical thinking was composed of three dimensions: a logical dimension, a critical dimension, and a pragmatic dimension. Ennis maintained that in applying any critical thinking skill, certain criteria or standards of that skill must be known.

The early emphasis on skill development was changed somewhat with the publication of Bloom's Taxonomy of Educational Objectives: Cognitive Domain (1956). Many social studies educators began to view critical thinking from a different perspective and endeavoured to use higher order questions to promote it.

The review of research studies conducted in this discipline reveals that they can essentially be divided into two sections: studies stressing the use of higher order questioning and studies relating to the teaching of specific skills or materials stressing such skills. This division reflects the two paradigm approach to critical thinking as viewed by social studies educators. Studies using higher order questioning are representative of the mental paradigm whereas the studies involving the teaching of specific skills or materials stressing such skills are indicative of the logical paradigm. The studies will

illustrate which of the two approaches has been most successful in the promotion of critical thinking in social studies.

Studies Utilizing Question Types

The effects of question types on the development of critical thinking was investigated by Hunkins (1970). His study involved two hundred and sixty, sixth grade students who were randomly assigned to one of two experimental groups. Group A received text type material stressing questions requiring analysis and evaluation as delineated by Bloom. Group B received text type materials which emphasized questions at the knowledge level. Pupils used the individually programmed instructional materials for 35 minutes per day for four weeks. The Social Studies Inference Test was used to measure changes in critical thinking ability. Analysis of covariance revealed that students who used only analysis and evaluation questions did not achieve significantly better than students who were exposed only to knowledge questions.

Similar results were obtained by Cohen (1973). His study examined classroom questions of tenth and eleventh grade Science, English, and Social Studies teachers to determine if frequency and type of questions were related to changes in critical thinking. The study involved 42 teachers and 263 students for a 20 week observation period. Critical thinking ability was measured by the

Cornell Critical Thinking Test, Level X and teacher questions were analyzed from audiotapes. Cohen concluded that no significant differences between cognitive level of questions existed between Science and non-Science teachers and that high level questions had no significant effect on student critical thinking ability.

Beseda (1973) also studied the effects of different levels of questioning, not only on critical thinking ability but on academic achievement as well. The experimental group consisted of 8 student teachers and 258 public school students. Teachers in this group received 12 hours of training in questioning techniques and received feedback each week from coded observations by the researcher. The teachers in the control group did not receive similar training in questioning techniques nor did they have the benefit of feedback. Students (N = 263) received regular instruction. Achievement was measured by the Iowa Tests of Educational Progress and the Sequential Test of Educational Progress. The Watson-Glaser Critical Thinking Appraisal tests were used to evaluate critical thinking ability.

Although the experimental group was exposed to more higher level questions, no significant improvement in academic achievement over the control group was detected. Contrary to the expected outcome, the control group scored significantly higher than the experimental group on gains in critical thinking ability. According to Beseda, the

high level of divergent questions acted as a deterrent to the development of critical thinking skills.

Caravan (1979) studied the effects of low and high level questions on critical thinking ability and the retention of gains in critical thinking ability effected by question types. High and low level questions were categorized according to Bloom's Taxonomy of Educational Objectives: Cognitive Domain. The study involved 142 tenth grade students who were randomly assigned for a three week period to one of the three groups: group one received low level questions; group two were taught by using high level questions; group three acted as a control group and were not subjected to any particular questioning approach.

Critical thinking ability was measured by the Watson-Glaser Critical Thinking Appraisal, Form YM. The assessment was given as a pre-test, post-test, and a delayed post-test administered one month after the termination of the experimental treatment. Analysis of variance revealed that the levels of questioning had no significant influence on critical thinking ability.

The fact that high level questions do not appear to produce significantly higher critical thinking abilities than low level questions may be due to the cognitive preference of students. Heller (1980) examined the cognitive preference of 250 seventh grade Social Studies students by using the Cognitive Preference Profile.

Independent variables such as text, teacher's cognitive preference, student achievement, and student sex were correlated with students' pre- and post- test scores for the Preference Profile. Results show there is no evidence to indicate that students prefer to learn critical thinking when exposed to curriculum materials and teachers which emphasize high level questions versus those who are exposed only to memory and application levels of questions. Teachers who preferred lower cognitive questions tended to produce similar preference in their students. High level achievers preferred the memory level of questions whereas the low achievers indicated preference for higher level questions. No sex differences in cognitive preference were detected.

Although Bloom's taxonomy has been widely endorsed by social studies educators, the finding of research studies tend not to support the effectiveness of this mental paradigm view of critical thinking. This approach appears to be unrealistic in that it maintains that students can acquire critical thinking skills by simply being exposed to higher order questions. In addition, it must be remembered that as Bloom himself acknowledges, the taxonomy was designed as a means of classifying educational objectives and not as a means of enhancing critical thinking skills (Bloom, 1986). Paul (1985a), Walsh and Paul (1986), and Wright and LaBar (1986) suggest that higher order questions can stimulate critical

thinking but first the criteria or standards for making judgments must be explicitly taught.

Studies Relating to Teaching Methods

Many studies have been conducted to determine if teaching specific skills significantly improves critical thinking ability. Henderson (1958) designed a study to test the hypothesis that knowledge of (1) determining the meaning of an expression, (2) deciding whether a statement is true or false, (3) deciding whether an argument is valid and (4) justifying opinions and evaluating other people's justification of their opinions, would lead to improvements in students' ability to think critically.

The experiment, involving 36 teachers and 1,500 students, was conducted over a two year period. Instructional materials were developed during the first year and the hypothesis tested the second year on grades 9-12 students taking English, Geometry, Science, and Social Studies. The investigator found that the experimental group, which used the prepared materials, made significantly greater gain from September to June on the Watson-Glaser Critical Thinking Appraisal Test than the control group. However, no significant differences between the two groups were detected when tested on the American Council on Education Test of Critical Thinking, nor were any significant differences detected on tests of academic achievement.

Selected students involved in this study were requested to participate in a "free response test involving evaluation of an argument in a fictitious letter to an editor". Students who had used the experimental material scored significantly higher than students to whom they were compared.

Shaver (1962) acknowledges that while results of the Henderson study appear to be contradictory it does give credence to the position that teaching the necessary specific concepts is the most effective means of instruction for critical thinking.

The teaching of specific critical thinking skills was also emphasized in a study conducted by Rothstein (1961). Two groups of eleventh grade History students, matched according to mental abilities, English reading, and critical thinking skills were selected for the study. The control group was instructed in the "conventional manner" whereas the experimental group received instruction with emphasis on the following critical thinking skills: comparing sources of various kinds; interpreting data, drawing inferences, and finding assumptions; identifying strong and weak arguments; evaluating thinking as to its relative criticalness or dogmatism; developing sensitivity to language and meaning and; augmenting student ability to draw conclusions from evidence and in differentiating fact from judgment.

Results of the thirty-five week study revealed that students who had received specific instruction in critical thinking skills achieved significantly higher scores on a critical thinking test than students taught by the regular or conventional methods. In addition, no significant correlation was found to exist between critical thinking ability and other school grades.

The influence of interaction between teaching method and student personality characteristics on the development of critical thinking skills was investigated by Shaver and Oliver (1968). The study, involving 125 seventh and eighth grade students, examined how various personality traits interacted with instruction involving the recitation method and socratic method. Three aspects of critical thinking - clarifying language, determining matter of fact and, making value claims, were emphasized in both groups. All students received background information to an issue in the same manner. Students in the socratic group were then encouraged to take and defend a stand on public issues whereas students in the recitation group experienced no personal discussion of the various issues. Results of the experiment demonstrate no significant differences between the two groups both in terms of social studies knowledge and general reasoning. Students in the socratic group did score significantly higher than the recitation group on tests developed by the

investigators to evaluate concepts covered by instruction. The researchers also conclude that no personality variables were found to interact consistently with the two styles of presentation.

Wright (1976) also investigated the interaction between two instructional methods and aptitude variables. The primary purpose of the study was to determine whether instruction that was based on a theoretical model of critical thinking would yield significant learning among elementary students. Two subtasks were to determine whether deductive expository or inductive discovery reasoning would be more beneficial and whether each method was appropriate for different groups of learners. The subjects, 369 sixth grade students, were randomly assigned to either the deductive expository or inductive discovery treatments. The treatment consisted of eight 40-minute periods of instruction in critical thinking. Critical thinking was envisioned as the application of such skills as discriminating between logical and illogical reasoning, identifying fallacies in illogical instances, discriminating between critical and uncritical responses to fallacies, and making critical responses to illogical reasoning, and using concepts such as generalization, analogy and inference. Analysis of student aptitude variables was achieved through the use of five different instruments.

The results of Wright's study suggested that while no significant achievement differences were discovered among the groups, the deductive expository method had a significantly better effect on pupil attitude than did the inductive discovery approach. The experiment also demonstrated that different approaches are more appropriate for different learners. The author suggests that the Progressive Matrices score can be used to determine the match between student and method of instruction.

Several studies have been conducted comparing the effect of different modes of instruction on the development of critical thinking ability. Cox (1963) compared the reflective and traditional approaches to instruction. Cox's critical thinking model consisted of the following categories: (1) orientation, (2) hypothesis formulation, (3) definition of terms, (4) exploration, (5) evidence, and (6) generalization. Method A (reflective model) made extensive use of open-ended discussions resulting in students being able to conceptualize hypotheses. In Method B (traditional approach), the emphasis was upon factual recall. Results of post-tests indicated no significant difference between the groups in terms of achievement or critical thinking abilities. However, analysis of classroom instruction tapes did not support this conclusion. The tape analysis tended to support the hypothesis that students using the

reflective model were more adapt at using the critical thinking model.

The traditional lecture-textbook method was also compared to the case study approach by Hunkins and Shapiro (1967) to determine which was more effective in promoting the development of critical thinking skills. The subjects, 54 children in two fifth grade classes, were randomly assigned to one of two groups. The experimental group (N=37) received instruction for 16 class periods and dealt with ten case studies covering economics, equality and citizenship, freedom of speech, worship, and privacy. The control group received regular lecture-textbook instruction. Analysis of pre- and post-test results indicates a statistically significant improvement for students using the case study approach. No such improvement was found to exist between the pre- and post-test scores for students exposed to the traditional lecture-textbook mode of instruction.

The traditional method of instruction was also compared to the cognitive approach by Cory (1975). The cognitive approach endeavoured to enhance critical thinking abilities through the use of analysis, synthesis, and evaluation levels of Bloom's Taxonomy of Educational Objectives: Cognitive Domain. Two secondary goals of the study were to ascertain whether high achieving students would make greater gains than low achieving students and whether males would make greater gains than females.

The study population consisted of 124 eighth grade students, equally divided into an experimental group and a control group. The necessary data was obtained by using the Watson-Glaser Critical Thinking Appraisal - YM, The Sequential Test of Educational Progress (STEP, Series 11 Form 3A - Social Studies), and a teacher constructed test. Post-test scores of standardized tests revealed no significant difference between the experimental and control groups. No significant difference was found to exist between male and female subjects. Only on the teacher constructed test did the experimental group score significantly higher than the control (traditional) group.

Rodhunsky (1968) used the Fenton Inquiry Sequence as a means of improving critical thinking in a grade ten History class. The random assignment of teachers established the three groups. Group one was instructed by teachers using the Fenton material. Group two had the same course objective without the materials and group three was taught in the 'regular' manner with the standard curriculum. Critical thinking ability was assessed on three separate occasions: pre-test, post-test, and delayed post-test. The delayed post-test was administered two weeks after termination of instruction. The Watson-Glaser Critical Thinking Appraisal forms were used as the instrument of evaluation. T-test analysis demonstrated that students using the Fenton materials developed significantly higher critical thinking ability than the

two other groups (P=.06). Results also show that the Fenton material proved most beneficial to those who had scored lowest on the pre-test.

Bate (1969) conducted a follow-up on the Rodnunsky investigation. His aim was to ascertain if increases in critical thinking abilities observed by Rodnunsky were maintained by students after a one year period. All three groups were again tested using the same instrument. Comparisons were made between the initial pre-test and delayed post-test of 1968 with scores obtained one year later. Results led Bate to conclude that students had indeed retained gains in critical thinking ability. Again, the trend observed by Rodnunsky, that the lowest in ability profited most by the Fenton material, was evident in the study by Bate.

A study designed by Pitts-Scarangello (1972) investigated the effects of an experimental text using built-in problem solving situations on critical thinking ability and achievement in fourth grade History. Five schools were randomly selected to participate in the study. Two classes, one experimental and one control, were established in each of the schools. The experimental groups (N=128) used a text designed to provide some thought provoking situation on each page. Students in the control group (N=126) used the regular text. In order to measure critical thinking and achievement gains, pre- and post-tests were administered to all students. Problem

solving ability (critical thinking) was measured by using an instrument designed by Ethel Maw entitled A Test of Critical Thinking for Grades IV, V, and VI. Achievement gains were measured by using The Delaware Test, designed by the researcher.

Analysis of results show that the mean gain scores of the experimental classes were greater than those of the control class on both critical thinking ability and achievement. However, only gains in critical thinking were at the significant level of acceptance. Correlation scores between the two test instruments in the experimental group was .98 whereas it was only .51 for the control group.

Curtis (1980) investigated critical thinking skills in non-academic social studies classes. The study involved 225 high school students who had previously been assigned to special classes for slow learners and non-achievers and non-academic vocationally oriented programs in eight schools, covering six school districts throughout British Columbia. Subjects, with ages ranging from 15 to 21, were randomly assigned to experimental and control groups.

The experimental group used the inquiry problem solving model for a four month period to investigate problems and issues of housing in a particular community. Students were exposed to the critical thinking skills as delineated by Ennis (1962) and were taught how to use

certain evaluation criteria for making judgments. Materials consisted primarily of a segment from Jack Webster's Open Line Show on housing problems as well as examination of a cartoonist's bias, point of view, ambiguity and contradiction from a comic book, entitled They Build Houses, Don't They. The control classes used the regular curriculum.

Critical thinking ability was evaluated by using the Cornell Critical Thinking Test as a pre- and post-test. Analysis of covariance revealed that the experimental curriculum had a positive significant impact upon the development of critical thinking skills.

An earlier study into developing critical thinking abilities of slow learners was conducted by Miller and Weston (1949). The aim of the experimental group in this study was to teach problem solving through the study and evaluation of different problems. Informal assessment by the researchers indicate an increase in critical thinking ability. However, when administered the Wrightstone Test of Critical Thinking in Social Studies, the students showed greater improvement than a control group only on the section dealing with the ability to draw conclusions. Shaver (1962) indicates that although the study lacks a strong scientific structure, it does indicate that slow learners can improve their critical thinking ability when specifically taught certain concepts that are deemed to be important.

It appears from these studies that an emphasis on specific skills and concepts tends to improve critical thinking ability. The results reinforce the belief that the logical paradigm appears to be most suitable for instruction in the social studies. Feely (1976) acknowledges that the logical paradigm is a "clarification rather than a break" with a past which stressed the development of specific skills. Today educators such as Beyer (1984b), Weddle (1986) and Wright and LaBar (1986) call for the development of curriculum materials based on Ennis' logical paradigm perspective of critical thinking. It appears, therefore, that the logical approach adopted for the present study can be justified in light of the evidence presented in this review of both the nature of critical thinking and critical thinking research in the social studies.

Summary

An examination of literature suggests that the concept of critical thinking does not lend itself to a single or simple definition. Many of the writings reflect the current debate among philosophers and educators as to the exact nature of the concept, how it should be taught, and to what extent a specific knowledge base is prerequisite for its development.

Since the early 1940's, many efforts, from various disciplines, have been expended on arriving at some

consensus as to the meaning of critical thinking. Proponents from these disciplines have delineated lists of skills deemed essential for the development of this skill. The Ennis paper of 1962 was a significant development because it had a major impact upon the perception of critical thinking by both philosophers and educators. For the first time, critical thinking was viewed as more than simply a list of various skills. Ennis proposed that critical thinking involved three dimensions. Today, in addition to the three dimensions, many educators agree that critical thinkers must also possess certain dispositions, sensitivities, and tendencies, referred to collectively as the critical spirit.

Studies involving the teaching of specific skills show significant improvement in the development of critical thinking as did instruction using the cognitive approach, case study approach, experimental text and the Fenton inquiry sequence. In addition, significant improvements were detected in studies involving slow learners. Some students were shown to be able to retain gains for at least a one year period.

The use of low and higher order questioning as an instructional approach produced contradictory results. While most studies showed little or no improvement in critical thinking abilities, one study indicated that high order questions proved detrimental to the development of these skills. It has been suggested that the cognitive

preferences of students may play an important role in these studies.

The review of social studies literature also reveals that initially the focus of critical thinking was in the teaching of specific skills. However, with the advent of Bloom's Taxonomy (1956) much of the effort was diverted into examining the effects of higher order questioning, as delineated by Bloom on the promotion of critical thinking ability. Recently, prominent social studies educators, such as Beyer (1984a), advocate a return to the teaching of specific critical thinking skills. Beyer (1984b) acknowledges that Ennis has provided the foundation from which these skills can be built.

CHAPTER III

INSTRUMENT AND METHODOLOGY

The purpose of this chapter is to provide a more comprehensive overview of the study and the procedures employed in conducting the research. It is sub-divided under the following headings:

(1) Selection of Instrument

(2) Population and Procedure

(3) Questions Examined and Statistical Procedures

Selection of Instrument

There are presently six commercially produced critical thinking tests which are readily available. In searching for a suitable instrument, these tests were examined and the characteristics of each are briefly presented below. Bearing in mind that the present study, using university students, attempts to measure critical thinking based on Ennis' definition which involves "reasonable reflective thinking that is focused on deciding what to believe or do", the limitations or suitability of each is noted.

Ross Test of Higher Cognitive Processes (1976a). The Ross test is designed to measure critical thinking ability in terms of the analysis, synthesis, and evaluation levels of Bloom's Taxonomy of Educational Objectives: Handbook 1.

The test, suitable for use in grades four to six, contains 105 items which are divided into eight parts and purports to measure such things as analogical reasoning, deductive reasoning, ability to identify missing premises, and ability to identify relevant and irrelevant information.

The instrument was not suitable for the present study for two reasons. First, the grade level of the test was inappropriate and secondly, the concept of critical thinking based on Bloom's Taxonomy differs substantially from that offered by Ennis and adopted for this study.

The New Jersey Test of Reasoning Skills. This test, developed by the Institute for the Advancement for Philosophy for Children (Shipman, 1983a), is designed for use with grades five through college level students. It contains 50 multiple choice questions covering some 23 skill areas involving reasoning, inquiry and concept formation.

The test, with a fifth grade reading level, is unsuitable for use with college level students. In fact, the manual warns that "some college students may object to its juvenile content". Since the test is curriculum specific and has a fifth grade reading level, its inappropriateness for this study is obvious.

Test on Appraising Observations. This aspect-specific critical thinking test (Norris and King, 1983) is designed to measure only one aspect of critical thinking - the ability to appraise observations. The test, suitable for use with high school and university level students, contains 50 items based on two stories that are presented. Part A presents a story involving a traffic accident whereas Part B deals with a story relating to the exploration of a river. In both sections, students are presented with a pair of statements and they must decide which is most believable in light of the evidence presented to that point.

Although based on Ennis' conceptualization of critical thinking, this test is aspect-specific and therefore unsuitable for this study which endeavours to measure several rather than one aspect of critical thinking.

Cornell Critical Thinking Test, Levels X and Z. The Cornell Tests, (Ennis and Millman, 1985) are based on Ennis' 1962 definition of critical thinking. Level X is designed for students from grades four to fourteen. Level Z is for use with a more advanced audience. Both levels attempt to measure aspects of critical thinking: such as inductive and deductive reasoning, observation ability, and ability to identify assumptions. These multiple-choice tests involve a series of questions in the context of a story.

These tests have several weaknesses which militate against their use for the present study. McPeck (1981) maintains that some items on the tests are measures of reading comprehension rather than critical thinking ability. A further weakness, according to Norris (1986) is that students using Level X with different experiences and sophistication "will tend to assume different things [from the story] and thus possibly justifiably choose different answers" (p. 137). These are weaknesses which generally characterize an objective type of critical thinking test. An essay test on the other hand, could compensate for these problems by encouraging students to supply reasons for their judgments. Information, derived from these reasons enables researchers to make inferences regarding the thinking processes employed by students. Such insight cannot be achieved by the use of objective types of critical thinking tests.

Watson-Glaser Critical Thinking Appraisal: Forms A and B. This test (Watson and Glaser, 1980a) is designed to measure many aspects of critical thinking. Both Form A and Form B evaluate the same aspects. Each contains 100 multiple-choice objectives and are sub-divided under the following subtests: inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments.

As with the Cornell Critical Thinking Tests some items on the Watson-Glaser are open to multiple inter-

pretations depending upon the assumptions made by the examinees (Ennis, 1984, Norris, 1986). Again, McPeck maintains that some items measure reading comprehension and not critical thinking ability. Because of these weaknesses, the instrument was judged unsuitable for the purposes of the present study.

Ennis-Weir Critical Thinking Essay Test: An Instrument for Testing and Teaching. This is the only available essay test (Ennis and Weir, 1985a) on the market. Its purpose is to evaluate processes of thinking rather than the product or conclusion emerging from that thinking, thus reducing many of the drawbacks associated with objective type tests. By examining the thinking process a more realistic appraisal of the critical thinking ability of students can be achieved. Through detailed written responses the different experiences and levels of sophistication of the subjects can often be detected. Such information becomes vital to the correct interpretation of test results. Without expensive and time consuming interviews with each subject, an essay test is the only means whereby the thinking process can be examined.

The test was selected for this study since it does measure the thinking process and is based on Ennis' conceptualization of critical thinking. A detailed description of the test and how it conforms to Ennis' perception of critical thinking follows.

The Ennis-Weir Critical Thinking Essay Test

The Ennis-Weir test attempts to measure many of the dispositions and skills outlined by Ennis in his goals for a critical thinking/reasoning curriculum (Appendix C). The five general categories of skills deemed essential by Ennis are outlined below.

I. Elementary Clarification
   - (1) Focusing on a question
   - (2) Analyzing arguments
   - (3) Asking and answering questions of clarification

II. Basic Support
   - (4) Judging the credibility of a source
   - (5) Observing and judging observation reports

III. Inference
   - (6) Deducing and judging deductions
   - (7) Inducing and judging inductions
   - (8) Making and judging value statements

IV. Advanced Clarification
   - (9) Defining terms and judging definitions
   - (10) Identifying assumptions

V. Strategy and Tactics
   - (11) Deciding on an action
   - (12) Interacting with others

Ennis has also provided a diagrammatical explanation of his critical thinking concept. Figure 1 illustrates how the goals outlined in Appendix C, and above, fit into the overall decision making process. The dispositions and abilities are depicted in the diagram by the Dispositions and Clarity boxes. Arrows flowing from these boxes indicate that the skills and dispositions are not confined

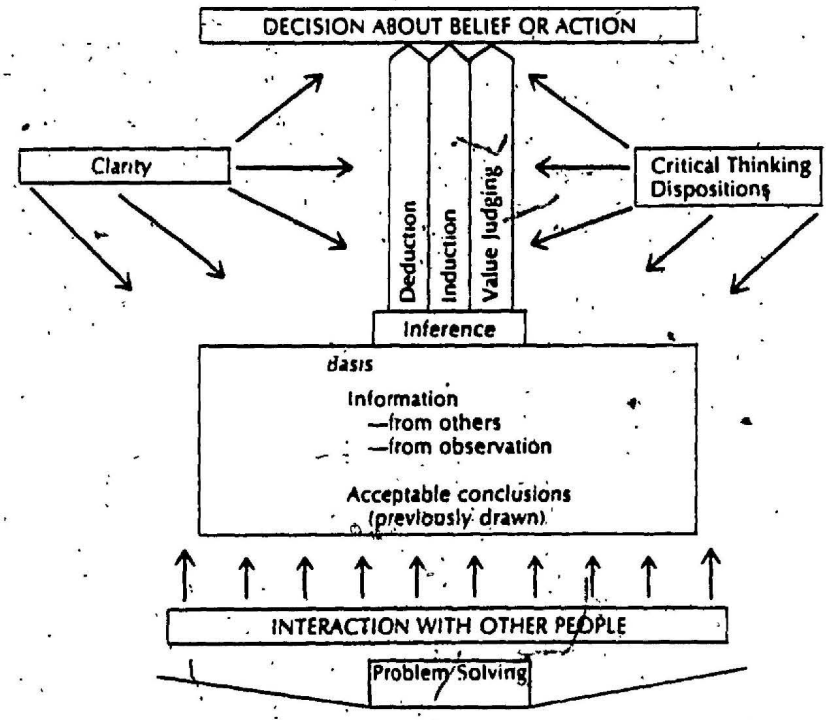


Figure 1: The Process of Deciding What to Believe or Do

Source: Ennis, R.H. (1985c) A logical basis for measuring critical thinking skills. Educational Leadership, 43, p. 47.

to a given area but rather are evident throughout the critical thinking process. The Basic Support box indicates that certain information and previously drawn conclusions must be taken into consideration before a decision can be reached. The Inference box, with arrows representing the three types emerging from it, bridges the gap between having the information (Basic Support) and making a decision. The two boxes at the bottom of the diagram, representing Strategy and Tactics, illustrate the connection between problem solving and interacting with other people to the overall decision making process.

The Ennis-Weir Critical Thinking Test (Appendix B) was developed to measure a variety of the dispositions and abilities as outlined by Ennis. According to the authors, the test purports to measure the following areas of critical thinking competence:

Getting the Point

Seeing the Reasons and Assumptions

Stating One's Point

Offering Good Reasons

Seeing Other Possibilities

(including other possible explanations)

Responding Appropriately to and/or Avoiding:

Equivocation

Irrelevance

Circularity

Reversal of an If-Then

(or other conditional) Relationships

The Straw Person Fallacy

Overgeneralization

Excessive Skepticism

Credibility Problems

The Use of Emotive Language to Persuade

A close examination of the test and the curriculum goals outlined by Ennis reveals that the test measures many of the dispositions and, to some extent, each of the five categories of skills as outlined by Ennis.

The authors have also attempted to eliminate some of the criticisms levied against multiple choice-objective type of critical thinking tests as noted earlier under the discussion of tests. In addition, the assumption is made by the authors that no specialized body of knowledge is required to perform the assignments required on the test. It is a "real world" test in that it is a task which every citizen should be capable of doing. Thus the requirement by Norris (1985b) that knowledge of the subject area in question is necessary before critical thinking can take place is satisfied.

The test, a letter to the editor of a fictional newspaper, provides students with an opportunity to evaluate both the form and content of another person's arguments presented in favor of prohibiting parking on all city streets between 2 a.m. and 6 a.m. The provided letter consists of eight numbered paragraphs. The

students are asked to evaluate and write a reply with nine numbered paragraphs. In the students' reply, the first eight paragraphs should be a paragraph by paragraph evaluation of the arguments presented by the writer. Students are requested to tell whether they believe the thinking good or bad. Reasons for their answers should also be provided. In the ninth paragraph, participants are asked to provide an overall evaluation of the total arguments presented. This newly designed test requires about 40 minutes to complete.

The test, developed for use with high school and university students, is accompanied by a scoring manual. The manual provides detailed directions as to what are acceptable and unacceptable responses to each paragraph. Indications are also given as to when the scorer should penalize responses that accuse the writer of faults he did not commit. Points awarded for the first eight paragraphs range from a -1 for an incorrect judgment to 3 for a complete answer. Scores on paragraph nine range from -1 to 5. The total test scores can range from -9 to 29.

After an evaluator has had sufficient experience, each test can be scored in less than 10 minutes.

Validity: Of the four types of validity: content, predictive, concurrent, and construct, the writers claim only content validity for the test at this time. This type of validity is applicable since the situation

presented would be familiar to all participants and would not require any technical skills or specialized body of knowledge. The authors also acknowledge that content validity is also claimed as a result of "judgment by experts in the field - after careful consideration" (Ennis and Weir, 1986b, p. 3).

Reliability: The reliability of the test is based upon the performance of 27 students in an introductory logic course at the university level and that of 28 gifted students in a grade eight English class. Two different markers were involved and obtained inter-rater reliabilities of .86 and .82.

The researcher acknowledges that the selected instrument is not without drawbacks. The fact that it is a new instrument that has not been extensively tested is an issue of concern. However, it is the only instrument available at the present time which suits the requirements of the present study. Boot strapping is inevitable considering the current state of the field. In fact, this study will provide an opportunity to evaluate the instrument itself.

Population and Procedure

Since the purpose of the present study is to determine the critical thinking ability of entering first year university students and leaving education students, it was necessary to select and evaluate the critical thinking ability of a random sample of first year students as well as the total population of selected high school methods courses offered by the Faculty of Education and taken by students in the fourth or fifth year of their degree program.

An attempt was made to obtain a 10 percent random sample of all first year students registered for English 1000. This course was selected on the basis that it is a required course for all students. Twelve classes were randomly selected for testing. However, due to time restraints and other circumstances beyond their control, four instructors were unable to grant permission for their class to participate. As a result, the final sample consisted of 164 first year English 1000 students, representing approximately six percent of the total population. The male/female ratio of the sample was 1:1.4 while the male/female ratio for the total population was 1:1.2. The ratio of the sexes indicates that the sample is fairly representative of the total population. In addition, an effort was made to ensure that selected English 1000 classes were offered by different instructors and scheduled in different time slots.

Senior students were selected from four high school methods courses offered by the Faculty of Education during the Winter Semester of 1986. The courses selected were as follows:

Education 4141	The Teaching of Literature in the Secondary School
Education 4163	Contemporary Approaches to the Teaching of Geometry in the Secondary School
Education 4181	The Development of Social Studies Curriculum in the Secondary School
Education 4201	Teaching Religious Studies in Secondary Schools

Students in Education 402X were also tested. This course, Internship in the Secondary School, contained students specializing in the areas listed above. The sample of senior students totaled 162 students. A summary of student participation is provided in Table 2.

Students do not join the education faculty until after their first year. Since the comparison in this study is between all first year students and students who finally enter education, there is a risk of bias from the chance that those students who enter education are not typical of those who enter the university as a whole.

Table 2

Level, Sex, and Number of Students

Who Participated in Study

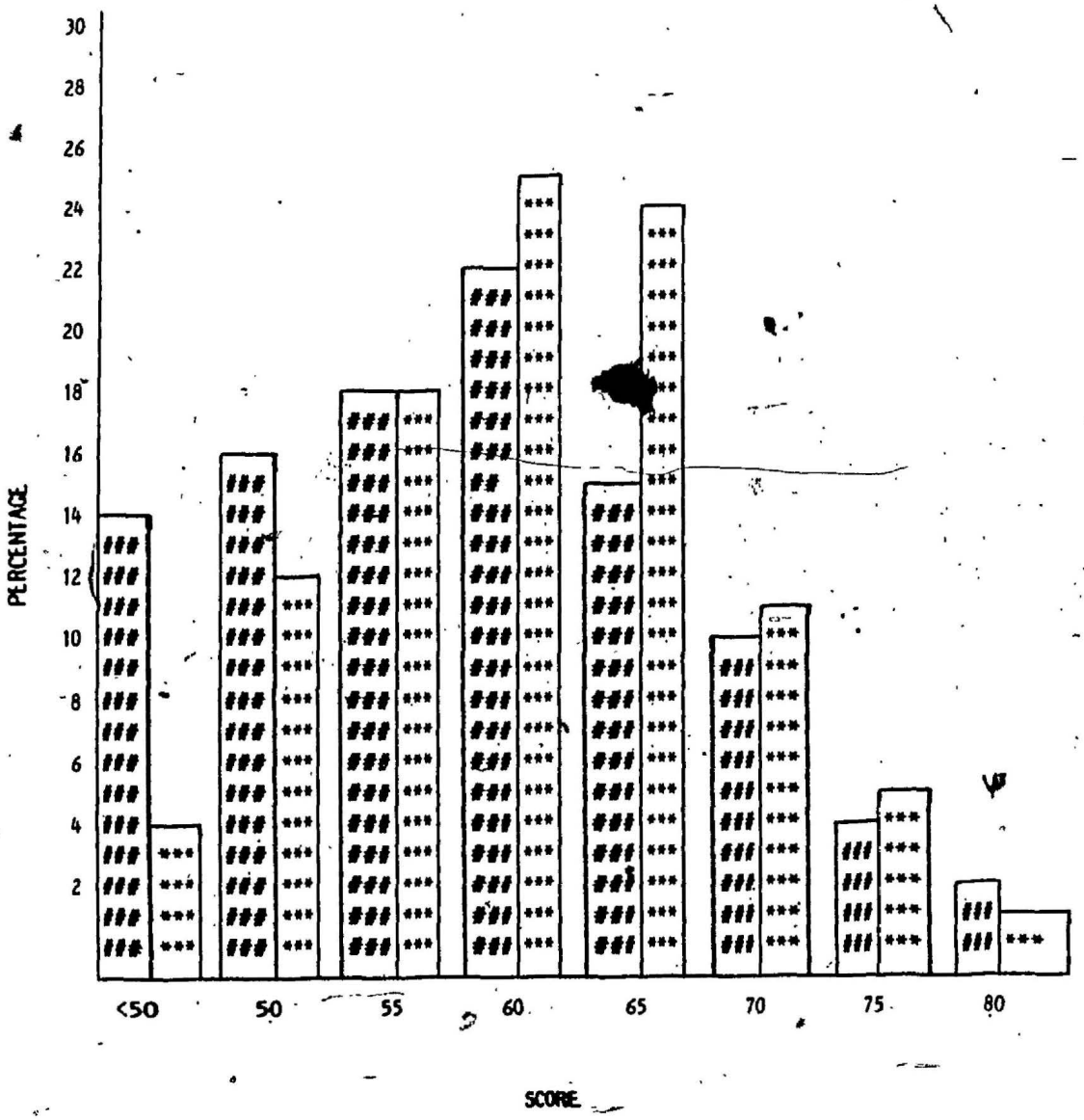
Level	Male	Female	Total
First Year	69	95	164
Fourth Year	32	18	50
Fifth Year or Greater	65	47	112

English 1000 grades were used to test for the existence of such bias. The grades of 2,457 students enrolled in English 1000 during the Fall of 1984 were examined and plotted on percentage bases as illustrated in Figure 2. The English 1000 grades of all students entering the Faculty of Education in the Fall, 1985, were also collated and plotted on a percentage basis as per Figure 2.

The greatest discrepancies exist in the range <50 and at the lowest B grade of 65. The first discrepancy can be explained by the fact that a large proportion of students scoring less than 50 percent on English 1000 may not continue at the University to enter any faculty. It is not clear why a second discrepancy exists, but overall the figure illustrates that there are no significant differences between scores of all students on English 1000 and those who were later admitted to the Faculty of

FIGURE 2

GRADE PERCENTAGES OF ENGLISH 1000



Grade Percentages of all English 1000 Students (Fall, 1984)

Grade Percentages of Students Admitted to the Faculty of Education on English 1000

Education. If the English 1000 score can be taken as an indication of academic ability, the assumption can be made that a typical education student is representative of the student body at Memorial University.

Inter-rater Reliabilities

After extensive study of the manual, a random sample of fifteen students' responses was selected for scoring. These were scored by Dr. Stephen P. Norris, a researcher in the critical thinking field, presently employed with the Institute of Educational Research and Development at Memorial University of Newfoundland, and by the author of the present study. The correlation between the raters' scores was .86.

When testing was completed, a random sample of thirty was selected. Again these were scored by Dr. Norris and the present author. In addition, the same samples were also sent to Dr. Robert Ennis, the co-author of the test with Dr. Eric Weir, at the University of Illinois. Due to work commitments, Dr. Ennis requested that Dr. Weir score the samples. The correlation between the three graders is as follows:

Rice - Norris correlation = .94

Rice - Weir correlation = .84

Norris - Weir correlation = .83

According to Coffman (1971) these are high inter-rater correlations for an essay test. These correlations

are as high as those reported by the test designers themselves. These high correlations, however, mask a scaling factor which will be discussed in Chapter 4 under the heading of concerns raised through the use of the Ennis-Weir Test.

Questions Examined and Statistical Procedure

The purpose of the present study is to examine the critical thinking ability of entering first year students and those nearing the end of their high school teacher preparation programme. To accomplish this goal, the following questions were examined:

1. Is there any difference in the critical thinking ability of first year males and first year females?
2. Is there any difference in the critical thinking ability of mature first year students (21 years of age or older) and other first year students?
3. Is there any difference in the critical thinking ability of urban first year students and rural first year students?
4. Is the critical thinking ability of senior students in the high school teaching program related to sex, year, number of philosophy courses, or area of study?

5. Is there any difference in the critical thinking ability of first year students and students in the fourth and fifth year of their Teacher Preparation Programme?

The first three questions, dealing with entering first year students, will be examined by using a three-way Analysis of Variance with total score on the Ennis-Weir test being the dependent variable and sex, maturity, and urbanness as the independent variables. Maturity is a term used to denote whether first year students are mature (21 years of age or older) or regular first year students. Some people might maintain that the critical thinking ability of mature first year students would be greater than that of other first year students since they would have been exposed to a wider variety of experiences. These experiences could influence the way they perceive the issue presented in the Ennis-Weir test.

The issue of urbanness was also explored with first year students. Urbanness refers to whether students are classified as urban or rural. Rural students would be those first year students coming from communities with a population of less than 1000 (Statistics Canada). This factor was examined since in Newfoundland there appears to be a widely held belief that the grade school education received by urban students is superior to that obtained by rural students.

Question four explored several variables which might be related to the critical thinking ability of students nearing the end of their high school teacher preparation programme. This question will be examined by a four-way Analysis of Variance. Total score will be the dependent variable and sex, year, number of philosophy courses, and area of study will comprise the independent variables. Sex was again examined to enable comparisons to be made with scores and gender of first year students. Distinction was made between fourth and fifth year students since, depending upon the number of courses taken, a student could be beginning a fourth year of study whereas another could be completing a fifth year of study. Such range in courses taken could conceivably influence critical thinking ability. The examination of senior students included only fourth and fifth year students since this is the level at which they would be required to register in appropriate instructional methods courses, thus distinguishing education students from students enrolled in other faculties.

The number of philosophy courses was also taken into consideration, for it is generally assumed that such courses tend to promote critical thinking. In addition, the Philosophy Department at Memorial offers a course specifically in critical thinking. It was important, therefore, to take the effects of Philosophy courses into consideration.

The final variable examined in question four was the students' area of study. It was important to study this variable since students from the areas of study selected often teach to some high scHoool social studies.

Urban/rural distinctions were not examined with the senior students because being exposed to a large university and city population for four or five years would probably erase any influence a rural environment would have had on critical thinking scores.

Question five dealt with the third purpose of this study: to make comparisons between the critical thinking ability of entering first year students and that of senior students. A two-way Analysis of Variance with total score being the dependent variable with sex and year acting as independent variables, will be used to examine this question.

CHAPTER IV

ANALYSIS AND PRESENTATION OF RESULTS

A detailed analysis of the collected data is presented in this chapter. The data will be examined according to the following categories:

(a) First Year Students, Questions One to Three

(b) Senior Students, Question Four

(c) Total Study Sample, Question Five

In addition, concerns raised through the use of the Ennis-Weir Test and several limitations to the study will be discussed.

Result Analysis of First Year Students: Questions One to Three

One of the purposes of the present study was to establish the level of critical thinking ability of first year students. A total of 164 first year students participated and the results of the Ennis-Weir Critical Thinking Essay Test indicate a mean score of 6.23 out of a possible score of 29. The standard deviation for the sample was 8.5.

Three questions were formulated to examine the sample of first year students and were examined using a three-way Analysis of Variance with the total score being the dependent variable and sex, urbanness, and maturity

acting as independent variables. The questions examined were:

1. Is there any difference in the critical thinking ability of first year males and first year females?
2. Is there any difference in the critical thinking ability of urban first year students and rural first year students?
3. Is there any difference in the critical thinking ability of mature first year students (21 years of age or older) and other first year students?

Table 3 provides a summary of the ANOVA results. An examination of the table reveals an interaction between sex and urbanness significant at the .07 level, and an interaction between sex and maturity significant at the .09 level. As a result of these significant interactions, unequivocal answers cannot be given to questions one to three which pertain to the main effects. The answer to question one is contingent upon whether students were urban or rural, mature or not. Similarly, the answers to question two and question three depend upon the sex of the subjects.

Table 3

Analysis of Variance Summary for First Year Students: Total Score by Sex, Urbanness, and Maturity

Source	df	MS	F	Sign. of F.
Main Effects				
Sex	1	215.1	3.1	.08
Urbanness	1	94.8	1.4	.25
Maturity	1	26.2	0.4	.54
Two Way Interaction				
Sex X Urbanness	1	228.8	3.3	.07
Sex X Maturity	1	199.2	2.8	.09
Urbanness X Maturity	1	2.6	0.0	.85
Three Way Interactions				
Sex X Urbanness X Maturity	1	36.0	0.5	.48
Residual	156	70.1		

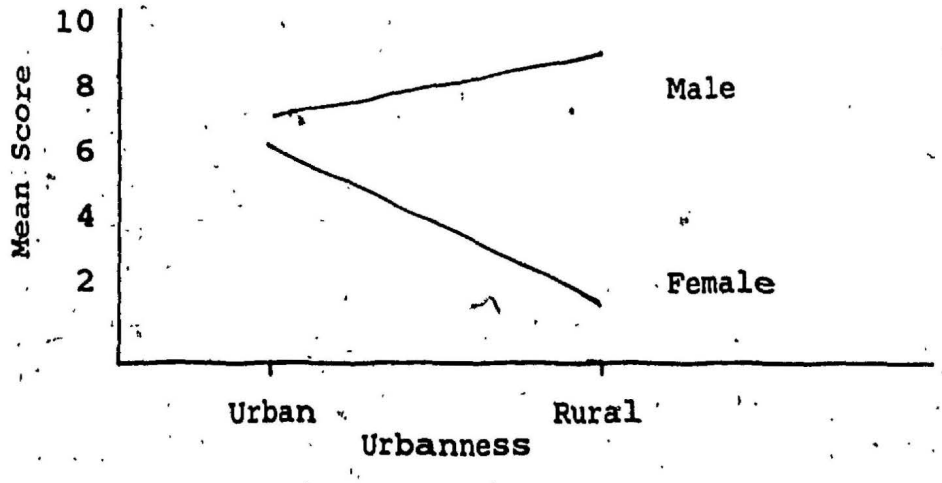


Figure 3: Mean Score of First Year Students by Sex and Urbanness

Figure 3 illustrates the interaction between sex and urbanness. The mean score for rural males (8.7) was higher than the mean score of urban males (7.3), which was the reverse of the trend of the female scores. Urban females had a mean score of 6.1 whereas rural females had a mean score of only 1.7. In both urban and rural samples, however, males obtained a higher mean score than females. Figure 3 also reveals that the difference in mean scores obtained between first year urban males and females is only 1.21 whereas the difference in mean scores between first year rural males and females is 6.92.

It is difficult to postulate why there was such a wide range of scores between first year rural males and females compared to the difference between urban males and urban females. Such considerations are beyond the scope of this study and in fact such questions were not realized until the present results were interpreted.

As stated previously, the interaction between sex and maturity was significant at the .09 level. Results indicate that 23 mature students involved in the study obtained a mean score of 7.43 with a standard deviation of 10.04. All other first year students (N = 141) achieved a mean score of 6.03 with a standard deviation of 8.15.

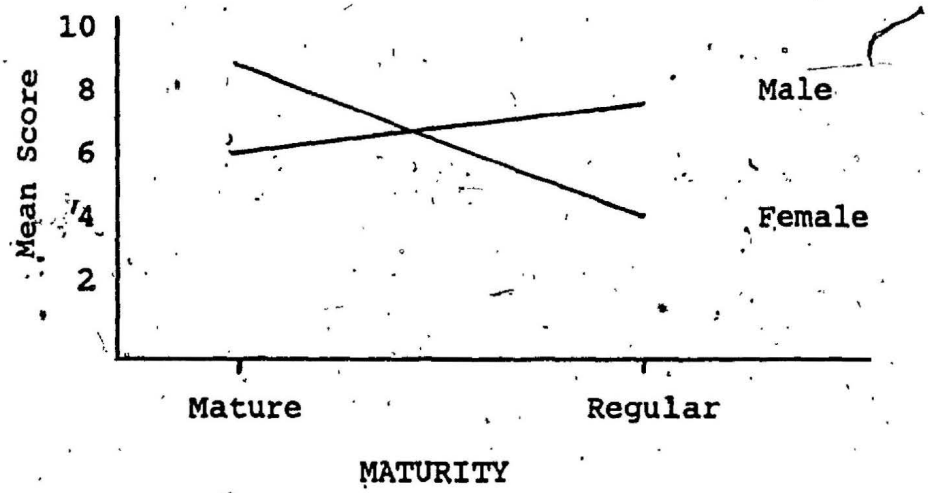


Figure 4: Mean Score of First Year Students by Sex and Maturity

The interaction between maturity and sex is depicted in Figure 4. It is evident that the mean score of mature females (9.0) was higher than the mean score obtained by mature males (6.0). This trend is the reverse of that obtained by regular students. The mean score of regular males (7.9) was higher than the mean score of 4.7 achieved by regular females.

It is interesting to note that the mean score of regular male students was higher than that of mature male students. The reasons for this trend is not evident from this study and is the opposite of that between mature and regular female students. The higher mature female score may be explained, in part, by the fact that only three of

the eleven mature females were from rural communities whereas 16 of 28 of the regular females were from rural communities. The low score obtained by rural females noted earlier may not have played a significant role in establishing the score of mature female students. Due to the small number of mature students involved, any further speculation would be untenable.

The examination of the interactions evident in the ANOVA results permit only qualified answers to be given to the questions about main effects. For instance, there was a difference in the critical thinking ability of first year males and first year females (Question 1) but only as it related to urbanness and maturity. To further investigate the interaction between sex and urbanness (Figure 3) and sex and maturity (Figure 4), an Analysis of Variance for Simple Main Effects was conducted. The results are presented in Table 4.

It is evident from the table and Figure 3 that rural males scored significantly higher ($p < .03$) than rural females. The table also reveals, in conjunction with Figure 4, that regular first year males (under 21 years of age) scored significantly higher ($p < .04$) than regular first year females.

Table 4

Analysis of Variance Table for Simple Main Effects: Sex at Urbanness, sex at Maturity

Source	SS	df	MS	F
Sex X Urbanness	228.8	1	228.8	3.3
Sex at Urban	51.94	1	51.94	.74
Sex at Rural	404.9	1	404.9	5.78**
Sex X Maturity	199.2	1	199.2	2.80
Sex at Mature	53.4	1	53.4	.76
Sex at Regular	344.8	1	344.8	4.91*
Sex	215.1	1	215.1	3.10
Within Cell	10935.6	156	70.1	

* $p < .04$

** $p < .03$

Result Analysis of Senior Students: Question Four

Question Four dealt with variables which might be related to performance scores by senior students. Results of the Ennis-Weir indicate mean scores of 10.84 and 10.29 for fourth and fifth year students respectively.

Question Four: Is the critical thinking ability of senior students in the high school teaching program related to sex, year, number of philosophy courses, or area of study?

This question was examined by using a four-way Analysis of Variance. Total score was the dependent variable and sex, year, number of philosophy courses, and area of study comprised the independent variables.

The Analysis of Variance summary is presented in Table 5.

Table 5 reveals that the only significant interaction emerging from the examination of this question is to be found between year and number of Philosophy courses. Figure 5 provides an illustration of this interaction which was significant at the .09 level.

Table 5

Analysis of Variance Summary for Senior Students: Total Score by Sex, Year, Number of Philosophy Courses, and Area of Study

Source	df	MS	F	Significance of F
Main Effect				
Sex	1	96.9	1.6	.21
Year	1	58.2	1.0	.33
Philosophy	1	168.0	2.8	.10
Area of Study	3	72.3	1.2	.32
Two Way Interaction				
Sex X Year	1	69.1	1.1	.29
Sex X Phil.	1	47.3	.8	.38
Sex X Area	3	53.5	.9	.45
Year X Phil.	1	177.5	2.9	.09
Year X Area	3	13.1	.2	.89
Phil. X Area	3	18.7	.3	.82
Residual	142	60.6		

The figure seems to illustrate contradictory trends. Fifth year students with two or more philosophy courses scored higher than fourth year students with the same number of courses whereas fourth year students with fewer than two philosophy courses obtained higher mean scores than did fifth year students who also had fewer than two

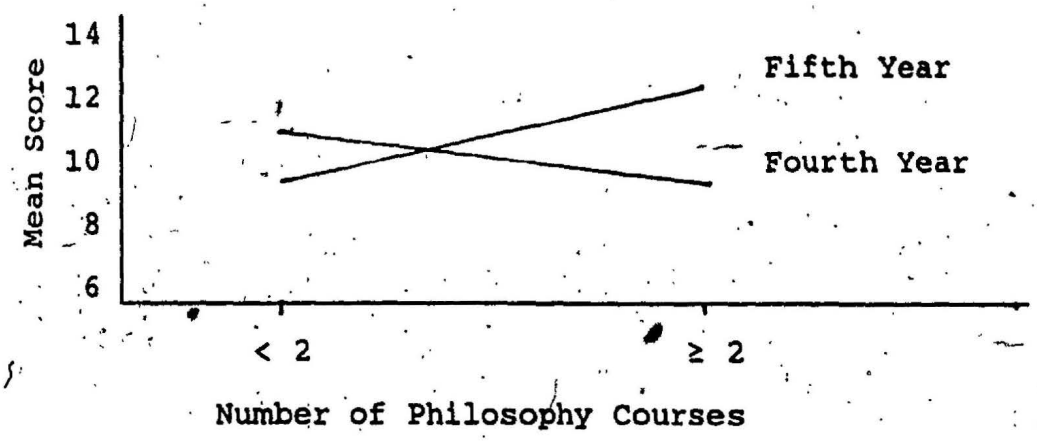


Figure 5. The Number of Philosophy Courses and Student Performance by Year

philosophy courses. In addition, fifth year students with two or more philosophy courses (N = 27) performed better than those with fewer than two philosophy courses (N = 85). On the other hand, fourth year students with two or more philosophy courses (N = 7) scored lower than those with fewer than two courses in philosophy (N = 42). The trend involving the fourth year students may be due to the fact that the seven students with two or more philosophy courses may not be a representative sample of all fourth year students at the university with two or more courses in philosophy. The type of philosophy course taken may also have some bearing upon the development of critical thinking skills. The type of Philosophy course taken is another variable worthy of further study.

The relationship between critical thinking ability and sex, year, and area of study did not produce any

significant interaction. However, Philosophy produced a significant main effect of .10, a fact which might be explained by the disposition to critical thinking of students selecting philosophy courses.

Result Analysis of Total-Study Sample: Question Five

Question Five was designed to examine the difference in critical thinking ability of first year and senior students.

Question Five: Is there any difference in the critical thinking ability of first year students and those in the fourth and fifth year of the high school teacher preparation programme?

Question Five was examined by using a two-way Analysis of Variance with total score being the dependent variable and sex and year acting as independent variables. The ANOVA summary is provided in Table 6.

It is evident from the table that no significant interactions exist between sex and year. The table also reveals that there was a significant main effect for both year and sex. These significant main effects now need to be investigated.

Table 6

Analysis of Variance Summary for All Students: Total Score by Sex and Year

Source	df	MS	F	Significance of F
Main Effect				
Sex	1	297.2	4.5	.035
Year	2	569.5	8.6	.000
Two Way Interaction				
Sex X Year	2	7.3	.1	.90
Residual	325	70.4		

Differences between First Year Students and Senior Students Enrolled in Four Major Areas of Study

In an effort to determine whether the significant differences in critical thinking between the Junior and Senior students was related to the four areas of study examined, an Analysis of Variance was performed on each grouping. A three-way analysis including sex, year, and area of study was not possible since the area of study variable was not applicable to first year students. The ANOVA summary tables are provided in Table 7 to Table 10.

Table 7

Analysis of Variance Summary: Total Score by Sex and Year for all First Year Students and Senior Students Majoring in English

Source	df	MS	F	Significance of F
Main Effect				
Sex	1	233.3	3.1	.070
Year	2	525.5	7.5	.001
Two Way Interaction				
Sex X Year	2	3.0	.04	.96
Residual	195	70.4		

Table 8

Analysis of Variance Summary: Total Score by Sex and Year for all First Year Students and Senior Students Majoring in Social Studies

Source	df	MS	F	Significance of F
Main Effects				
Sex	1	156.5	2.3	.130
Year	2	282.2	4.2	.017
Two Way Interaction				
Sex X Year	2	53	.8	.454
Residual	220	67.8		

Table 9

Analysis of Variance Summary: Total Score by Sex and Year for all First Year Students and Senior Students Majoring in Science

Source	df	MS	F	Significance of F
Main Effects				
Sex	1	335.0	4.8	.03
Year	2	176.8	2.5	.08
Two Way Interactions				
Sex X Year	2	66.3	.95	.39
Residual	195	69.8		

Table 10

Analysis of Variance Summary: Total Score by Sex and Year for all First Year Students and Senior Students Majoring in Religious Studies

Source	df	MS	F	Significance of F
Main Effects				
Sex	1	275.0	3.9	.06
Year	2	69.3	1.0	.37
Two Way Interactions				
Sex X Year	2	8.7	.1	.89
Residual	184	69.6		

The tables reveal that no significant interactions between sex and year occurred within either of the groupings. However, significant main effect differences for year were found to exist between first year students and senior students whose areas of concentration were English ($p = .001$), social studies ($p = .017$), and science ($p = .08$), but not religious studies.

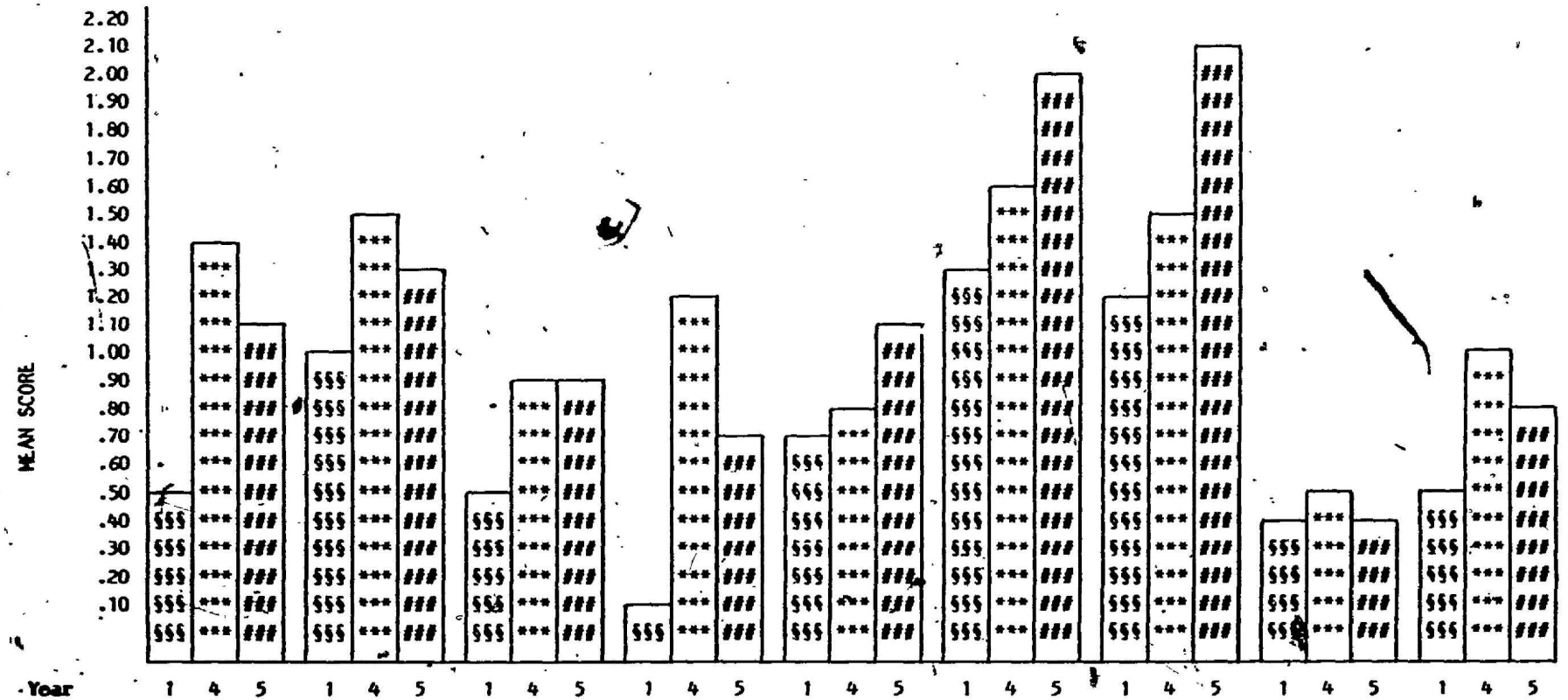
Significant main effect differences for sex were found to exist between first year students and senior students majoring in English ($p = .07$), science ($p = .03$), and religious studies ($p = .06$). No main effect significant differences for sex were detected between first year students and social studies majors.

Student Performance on Each Paragraph of the Ennis-Weir Test

Once the significant differences between first year students and senior students in the four areas of study were examined, Analysis of Variance was used to assess student performance on each of the paragraphs of the letter contained in the Ennis-Weir Critical Thinking Test. The results are presented in Figure 6. Such an analysis permits an evaluation of student performance by year on various aspects of critical thinking measured by the test.

FIGURE 6

MEAN SCORES OF ALL STUDENTS ON THE ENNIS-WEIR CRITICAL THINKING ESSAY TEST



Comparisons between first year and senior students could reveal the influence of university training on specific aspects of critical thinking.

It can be observed that senior students outperformed first year students on all nine paragraphs. However, the difference between mean scores was only significant for four of these paragraphs: Paragraph One ($p < .001$), Paragraph Four ($p < .001$), Paragraph Six ($p = .01$), and Paragraph Seven ($p < .001$). It can be concluded therefore that senior students are significantly better than first year students at:

(1) Recognition of misuse of analogy, and/or recognition of shift in meaning, and/or claim that incorrect definition has been stipulated.

(2) Recognition of circularity, and/or recognition that no reason is offered.

(3) Recognition of lack of controls in an experimental setting, and/or inadequate sampling, and/or "only one case", and/or "post hoc fallacy".

(4) Recognition of winning argument by definition, and/or recognition that a word has been made useless for empirical assertion, and/or claim

> that an incorrect definition has been asserted. (Ennis-Weir, 1986b, p. 14)

According to the goals established by Ennis, it appears that first year students are weakest at using Advanced Clarification and Strategy and Tactics (Appendix C).

An examination of what is being evaluated in paragraphs one, four, six, and seven may give some indication as to why senior students scored significantly higher than first year students.

Paragraphs one and seven, according to the criteria and scoring sheet, measure a student's ability to deal with incorrect use of definitions. Paragraph four, although it evaluates ability to detect circularity, is also similar in that students are penalized for interpreting "undesirable" as "not desired". Again, application of definition is involved.

It appears from the results that first year students are unable to handle higher abstract mental operations such as evaluating inappropriate use of definition. They seem to be either unwilling or unable to challenge definitions established by others.

As indicated, senior students also scored significantly higher than first year students on paragraph six. This may be due, in part, to the emphasis placed on certain words in the paragraph. The same argument also

holds true for paragraph seven. In paragraph six the words "not one accident" and the sentence "Conditions are not safe if there's even the slightest possible chance of an accident" in paragraph seven, are placed in bold type. No reasons are provided to indicate why these are the only two sections of the letter presented in such a type. It may have been done to make the errors in the paragraph more obvious to the students. However, most students entering university have used the technique of underlining or highlighting important things in their texts for examinations which often require the recall of specific factual content. The texts used in high schools often use bold or italicized type to illustrate important information. Students may have interpreted the bold type as being the most important and regarded it as a definitive statement which should not be questioned. If this is so, they probably would have disregarded the first part of the sentence in paragraph six which specified that the experiment was for only one four hour period and concentrated instead on the "not one accident" claim. In paragraph seven, simply believing what was in the bold type would automatically lead to an incorrect conclusion.

Figure 6 also illustrates the fact that fourth year students performed better than other students on five of the nine paragraphs. However, none of the observable differences were statistically significant.

Fifth year students achieved a higher mean score than all other students on three of the nine paragraphs. They performed significantly better ($p = .05$) than fourth year students only on Paragraph Seven which measures the ability to recognize unjustified attempts to win arguments by definition.

Concerns Raised Through the Use of the Ennis-Weir Test

The Ennis-Weir test is a newly designed critical thinking test that has not been used extensively. Since it is an essay test, unlike other critical thinking tests, an evaluation of its use serves an important function. The issues raised may have implications for its acceptability as a valid indicator of critical thinking ability. The concerns raised through its use in the present study are addressed below.

(1). Inadequacy of Test Directions. The overall intent of the Ennis-Weir test of critical thinking is to evaluate the writer's thinking process as evident by the arguments presented rather than the product or stance taken by the writer. However, the directions provided do not only lend themselves to this interpretation. The directions state that:

> ...For each paragraph in the letter you are about to read, write a paragraph in reply telling whether you believe the thinking good

> or bad. Also write a closing paragraph about the total argument. Defend your judgments with reasons.
>
> Your answer should have nine numbered paragraphs. Numbers one through eight should give your reactions to paragraphs one through eight in the letter. Your paragraph nine should give your overall evaluation of the letter considered as one total argument. Each paragraph, including the last, should contain your reason(s).

Students are reminded twice more in the directions to provide reasons in each paragraph. However, when asked if the thinking was good or bad, many students seemed to interpret this to mean "Was the writer's conclusion good or bad?". As a result, in order to defend their judgments, as directed, students began to take positions and offer counter arguments on the issue of prohibiting overnight parking. Although participants are exhorted to supply reasons in each paragraph, it is never linked directly to the issue of whether it is the reasoning or the conclusion of the reasoning that should be evaluated. Many of the students tested by Norris and Thomey (cited in Norris, 1986) interpreted the test directions in the same manner. According to Norris (1986),

> to request an evaluation of someone's thinking is ambiguous between requesting an evaluation

> of the process of that thinking and requesting an evaluation of its conclusion, between evaluating how someone was thinking and evaluating what someone was thinking. (p. 143)

As Norris points out, this defect has implications for the validity of the test. If it does not measure what it is supposed to measure, a student's score may be indicative of something completely different from what the authors had planned.

(2) <u>The Possibility of a Cultural Bias.</u> Some students (6.7%) had problems in identifying with the situation presented in paragraph three.

> Paragraph Three: Traffic on some streets is also bad in the morning when factory workers are on their way to the 6 a.m. shift. If there were no cars parked on these streets between 2 a.m. and 6 a.m., then there would be more room for this traffic.

Many students doubted the writer when he claimed that there was heavy traffic on the street at this hour of the morning. It was their opinion that most people commence work at 8 a.m. or 9 a.m.. This, no doubt, is due to the fact that most workers in Newfoundland begin their

workday at these times. If answers supplied by respondents are influenced by their cultural background the validity of the test comes into question. Adaptations to the test may have to be made to help prevent the possibility of cultural bias.

(3) Defects in the Scoring Manual. While Ennis and Weir have provided a fairly good scoring manual, there appears to be some inadequacy in scoring directions with regard to awarding points for certain objections. For example, in paragraph three full points are awarded if students recognize that the arguments presented are valid. If they failed to neglect mentioning that the argument applies to only busy streets, responses are not to be penalized. On the other hand, if the thinking of students were sophisticated enough to detect that the arguments are really invalid since they apply only to busy streets and not to all streets, no marks could be awarded if scoring directions were followed exactly.

Paragraph five also serves as a case in point.

Paragraph Five: If parking is prohibited from 2 a.m. to 6 a.m., then accidents between parked and moving vehicles will be nearly eliminated during this period. All intelligent citizens would regard the near elimination of accidents in any period as highly

desirable. So, we should be in favor of prohibiting parking from 2 a.m. to 6 a.m..

According to the scoring manual, there are three types of defects with this argument:

(i) The type of accidents that would be eliminated if the recommendations were adopted is a very special and restricted type - accidents between parked and moving vehicles.

(ii) Other things such as inconvenience and economic cost to residents might be judged more important than eliminating accidents between parked and moving vehicles.

(iii) The argument describes only one possible way of eliminating accidents of this type.

Three points should be awarded if respondents indicated that inconvenience and economic cost to the residents could be substantial if they are not allowed to park on city streets overnight. If, however, students mention the issue of inconvenience and economic cost in other paragraphs, especially paragraphs one and four, as a rebuttal to the arguments presented, no points can be awarded.

(4) Scaling factor resulting from mean differences. To establish inter-rater reliabilities, three readers scored

30 randomly selected samples. As mentioned previously, the obtained inter-rater reliabilities were high for an essay test. This simply means that the three scores were fairly consistent in ranking students' performances. These high reliabilities, however, mask differences in obtained means. These differences tend to reduce the acceptability of any absolute, as opposed to relative, critical thinking levels established. Table 11 illustrates the mean scores and standard deviations obtained from the three readers.

Table 11

Mean Scores and Standard Deviations of Inter-rater Scores

Scorer	1	2	3
Mean	4.6	8.3	14.9
S.D.	7.1	9.6	9.3

1. Dr. Stephen P. Norris, Institute for Educational Research and Development, Memorial University
2. Present Researcher
3. Dr. Eric Weir, Co-author of Ennis-Weir Test, University of Illinois

It is evident from the table that with a 10.3 point range in the mean, it is difficult to make any definitive statement with regards to actual thinking levels. This is reflected in Table 12 which projects the scores of students at Memorial obtained by this researcher to scores which might have been obtained had another rater scored their tests.

Table 12

Mean Scores of Students at Memorial University Projected According to Means Established By Inter-rater Scores

Group	1	2	3
Present Researcher	6.23	10.54	10.3
Norris	2.53	6.84	6.6
Weir	12.83	17.14	16.9

1. First year students enrolled in English 1000.
2. Fourth year Education students.
3. Fifth year Education students.

An examination of the table reveals the substantial differences in means which might be expected by the three readers. Such a wide range of means can change how one would interpret the results. For example, the mean scores obtained from Norris would indicate very poor student performance whereas the mean scores assigned by Weir would

place student critical thinking ability in a much more favourable light. This scaling factor thus tends to limit the use of the test. It can serve as a teaching instrument and as a means of ranking students within a particular group(s) that are tested by the same researcher. Any other comparisons using total means might be invalid.

Limitations of the Study

1. The assumption is made that in cases where individual instructors administered the test they followed the specified directions. Instructors were requested to simply distribute the test to their classes and let students follow the clearly specified direction. It must be assumed that instructors did not make any comments which would affect students' perceptions and thus their answers to the test.

2. Limited access to sample. An effort was made to obtain a ten percent random sample of first year students enrolled in English 1000. Due to difficulties, noted in Chapter 3, this goal was not accomplished. Instead, a random sample of six percent was achieved. However, this group representing 164 first year students was fairly

representative of the male/female ratio at the university.

3. There are limitations of the use of the instrument as discussed in the previous section.

The first two limitations concerning instructors' following directions and access to sample would, in most likelihood, not significantly affect the findings of this study. However, the limitations of the instrument itself, as previously discussed, provide an issue of concern and one which must be reflected in the interpretation of results. When issues involving validity, cultural bias, defects in scoring manual, and the scaling factor are taken into account, the results must be viewed with caution. It must be remembered, however, that these factors would affect the score obtained by particular groups rather than significantly influence the trends observed within these groups. The observable trends within groups would probably not be significantly affected by the variables listed since all participants would, to some degree, be exposed to the same deficiencies of the test. The variables, however, become significant when trying to assign definitive levels of critical thinking or when comparisons are made between groups from different cultures and scored by different researchers.

CHAPTER 5

SUMMARY AND DISCUSSION OF EDUCATIONAL IMPLICATIONS

Summary

The present study was established to investigate:

(1) the critical thinking competence of beginning high school social studies teachers and explore several factors which might be related to that level of competence. Since, in Newfoundland, social studies is often taught by graduates from other disciplines it was necessary to evaluate the critical thinking ability of students enrolled in a variety of specializations. Beginning teachers were defined as those university students who were nearing the end of their high school teacher preparation programme, and enrolled in appropriate methods courses.

(2) the critical thinking ability of first year students at Memorial University and examine factors which might contribute to established levels. First year students were tested, in part, to give some indication of the success of the current high school program, which includes a core of at least four social studies courses, in the promotion of critical thinking skills.

(3) the relationship betweem university training and critical thinking ability. Comparisons were made between first year and senior students, not only to detect

difference between the two groups but also to explore where these differences occurred.

More specifically, the study attempted to answer the following questions:

1. Is there any difference in the critical thinking ability of first year males and first year females?
2. Is there any difference in the critical thinking ability of mature first year students (21 years of age or older) and other first year students?
3. Is there any difference in the critical thinking ability of urban first year students and rural first year students?
4. Is the critical thinking ability of senior students in the high school teaching program related to sex, year, number of philosophy courses, or are of study?
5. Is there any difference in the critical thinking ability of first year students and students in the fourth and fifth year of their high school teacher preparation programme?

Instrument

The Ennis-Weir Critical Thinking Essay Test: An Instrument for Testing and Teaching, was used to evaluate the level of critical thinking of entering first year students and that of fourth and fifth year students

nearing the completion of their high school teacher preparation program. The test involves an evaluation of a letter written to the editor of a fictitious newspaper which presented arguments to prohibit overnight parking between 2 a.m. and 6 a.m. on all city streets.

The test is designed to measure a person's critical thinking ability in the context of argumentation. The aim is to have subjects evaluate the manner in which arguments are presented thus allowing researchers to evaluate the thinking process of these students rather than the products of their reasoning. This is a newly designed test and is based on Ennis' conceptualization of critical thinking and it measures many of the goals outlined in his critical thinking/reasoning curriculum (Appendix C).

Procedure

First year students participating in the study were tested in classes which were randomly selected from all English 1000 classes offered during the Fall and Winter semesters 1985-86. The sample represented students registered in courses offered by different instructors and in different time slots. One hundred and sixty-four first year students participated.

The total population of senior students registered in four high school methods courses, representing different areas of specialization, and students enrolled in a course entitled "Internship in Secondary Schools" were surveyed.

A total of 50 fourth year and 112 fifth-year students took part in the study.

Grading of the essay tests was done solely by the researcher. However, inter-rater reliabilities were established with Dr. Stephen P. Norris, Institute for Educational Research and Development at Memorial University of Newfoundland and Dr. Eric Weir, co-designer of the instrument, University of Illinois.

Findings

Analysis of performance by first year students indicated a mean score of 6.23 out of a possible score of 29. The ANOVA results revealed an interaction between sex and urbanness significant at the .07 level and an interaction between sex and maturity significant at the .09 level. As a result, Analysis of Variance for simple main effects was conducted for sex at urbanness and sex at maturity. The results of this investigation revealed that the score of rural males was significantly better ($p < .03$) than the score of rural females. In addition, regular males achieved significantly higher ($p < .04$) scores than did regular females.

The ANOVA results of senior students indicated that no significant differences were detected among students enrolled in the high school teacher preparation programme according to sex, year or area of specialization. However, significant interaction, at the .09 level, was

detected between year and the number of philosophy courses. This interaction, however, produced puzzling trends. Fifth year students with two or more philosophy courses scored higher than fourth year students with the same number of courses whereas fourth year students with fewer than two philosophy courses obtained higher mean scores than did fifth year students with fewer than two philosophy courses. In addition, fifth year students with two or more philosophy courses performed better than those with fewer than two philosophy courses. The opposite trend was true for fourth year students.

When comparisons were made between the total scores obtained by first year students and those achieved by fourth and fifth year students, significant differences were detected. This finding is consistent with a longitudinal study by Lehman (1963) who found significant changes in critical thinking ability from freshman to senior years. An instrument entitled Test of Critical Thinking (American Council of Education 1953) was used by Lehman as the evaluation instrument.

Upon further analysis of the comparisons between junior and senior students in the present study, it became evident that the significant differences exist only between first year students and those senior students specializing in English, social studies, or science. While the senior students, as a group, scored higher than first year students on all paragraphs of the Ennis-Weir

test, the difference in scores was only significant for four of the nine paragraphs. Fifth year students scored significantly higher than fourth year students only on paragraph seven.

An important outcome of the present study was a fairly extensive evaluation of the instrument, the Ennis-Weir Critical Thinking Essay Test. Several questions were raised about the use of this test as a valid measure of critical thinking ability. It was noted that possible problems with test directions and a possible cultural bias could influence the validity of the test. Possible deficiencies in the scoring manual were also addressed. Another limitation dealt with the scaling factor detected between different readers. Three readers scored a random sample of 30 tests and despite very high correlations, an examination of the means of the samples revealed a 10.3 point range among readers. As a result, an attempt to establish actual levels of critical thinking ability was frustrated. Whether this is a problem with the test or with the particular combination of raters in this study is not known.

If these possible problems with the test are verified in other studies most could be rectified by making minor changes in the test, test directions, and scoring manual.

Even with no changes the test can still serve as a valuable teaching and testing instrument for researchers interested in detecting flaws in reasoning and ranking

students' performance within a particular group. The test, based on a comprehensive model of critical thinking, is the only instrument available, apart from personal interviews, which permits an evaluation of the thinking process of participants rather than simply an evaluation of the products of the thinking process. This fact alone speaks highly of its value as a diagnostic tool.

Discussions of Educational Implications

Even when the scaling factor discussed in Chapter Four and above is taken into consideration, the results of the present study are less than encouraging. Even with the most generous rater the average score for the reliability sample would be just over 50% of the total score possible. The study appears to reinforce what researchers and educators have been saying for some time: students are taught what to think rather than how to think. According to Olsen, the present system of education

> ...tends to teach us to conform; to solve non-creative problems (those with the answers at the back of the book); to reward coming up with an idea and not taking the action required to implement our idea. It makes us trust written material such as books too much; leads us to believe that others who are more wise have the

> real answers and separates learning from doing ... our educational system leads us to believe that failure is wrong and of no value. (cited in Walsh and Paul, 1986, p. 10)

Even if Olsen overstated the problem, the implications are clear. The primary purpose of schools, which is to educate youth on how to function to the best of their ability in a democratic society, will not be realized. It is evident from the present study that, if the Ennis-Weir test results are to be taken seriously the instructional techniques used at the university level as well as the high school level do not promote the development of critical thinking skills to the extent that many people desire. Although results indicate that senior students, in some areas of specialization, performed significantly better than first year students, the mean scores indicate that prospective teachers are not well qualified to instruct others in this area. This finding could have serious implications for the educational system in Newfoundland.

If students who are completing high school and entering university do not have a high level of critical thinking ability and prospective teachers who are about to enter the teaching profession have not mastered the skills sufficiently to instruct others, the cycle becomes self-perpetuating. In examining the critical thinking competence of beginning social studies teachers, graduates

from other academic areas were also evaluated and no significant differences, based on total score, were detected among the groups. This suggests that although the Aims of Public Education for Newfoundland and Labrador, the Master Guide for Social Studies and various other curriculum guides constantly reiterate the importance of critical thinking to the school curricula, in all likelihood, the skills will not be taught systematically in the classroom.

Educators must become more cognizant of the discrepancy between the stated goals of critical thinking and lack of evidence regarding actual instruction. The public of Newfoundland, as evident by the Warren (1978, 1983) studies expect more in the way of critical thinking instruction than the students are presently receiving. In Chapter One of the present study several factors were acknowledged as reasons why little critical thinking instruction takes place within the classroom. Chief among these was the fact that educators did not have a clear conception in their own mind of the exact meaning of critical thinking. Today, however, educators from many disciplines appear to accept the definition of critical thinking as envisioned by Ennis. His comprehensive description of critical thinking can serve as a foundation, leaving the onus on those who control the educational process to determine the priority of critical thinking in the curriculum.

If critical thinking skills are essential for effective citizenship, as educators and theorists have postulated, the status quo can no longer be viewed as acceptable. However, as Walsh and Paul (1986) indicate, the needed reform will occur only if people are willing to take bold, decisive and immediate action. They, like Bereiter (1984) and others, believe that if the exercise of critical thinking skills should form an integral part of our daily decisions, then critical thinking should not be taught as a separate course or enrichment but rather as an integral component of the entire school curriculum. Walsh and Paul have outlined what they consider to be factors essential for effective integration of critical thinking into the curriculum, thus moving critical thinking from an ideal into a reality (see Appendix D).

While there continues to be a debate over how best to teach critical thinking skills the results of the present study may be an issue of concern for the University and the Department of Education. Simply being aware of the findings could provide an incentive for re-evaluation of the current instructional techniques. In fact, it may require self-evaluation on part of educators. Often, without being aware, many teachers place too much emphasis on content coverage, standardized test performance and unquestioning acceptance and docility in students (Walsh and Paul, 1986; Beyer, 1984a, Olsen, 1984).

The hope is that the results obtained and the issues raised in this study will focus attention and stimulate debate among educators which could eventually yield further dividends for educational theory and practice.

BIBLIOGRAPHY

BIBLIOGRAPHY

Aims of public education in Newfoundland and Labrador (1984) Province of Newfoundland and Labrador.

Allen, R.R. and Rott, R.K. (1969) The nature of critical thinking. Report from the Concepts in Verbal Argument Project. Theoretical Paper No. 20, University of Wisconsin. Research and Development Center for Cognitive Learning. Madison, Wisconsin. ERIC Document ED 036 861.

Anderson, H. (Ed.) (1942) Teaching critical thinking in social studies: 13th yearbook. Washington: National Council for the Social Studies.

Aylesworth, T.G. and Reagan, G.M. (1969). Teaching for thinking. New York: Doubleday and Co.

Bate, J.P. (1969) Retention of gains in critical thinking effected by a Fenton inquiry sequence. Unpublished Master's Thesis, University of Calgary, Alberta.

Bereiter, C. (1984) How to keep thinking skills from going the way of all frills. Educational Leadership, 42, 75-77.

Beseda, C.G. (1973) Levels of questioning used by student teachers and its effect on pupil achievement and critical thinking ability. (Doctoral Dissertation, North Texas State University, 1972) Dissertation Abstracts International, 32, 4214A.

Beyer, B.K. (1977) Back-to-basics in social studies. Boulder, Colorado: Social Science Education Consortium, Inc.

Beyer, B.K. (1983) Common sense about teaching thinking skills, Educational Leadership, 41(3), 44-49.

Beyer, B.K. (1984a) Improving thinking skills - defining the problem. Phi Delta Kappan, 65, 486-490.

Beyer, B.K. (1984b) Improving thinking skills - practical approaches. Phi Delta Kappan, 65, 556-560.

Beyer, B.K. (1985a) Critical thinking: What is it? Social Education, 49, 270-276.

Beyer, B.K. (1985b) Teaching critical thinking: A direct approach. Social Education, 49, 297-303.

Black, M. (1953) Critical thinking: An introduction to logic and scientific method. (2nd ed.) New York: Prentice-Hall.

Blair, J.A. (1986) The Keegstra affair. The History and Social Science Teacher, 21, 158-164.

Bloom, B.S. (Ed.) (1956) Taxonomy of educational objectives, handbook 1: Cognitive domain. New York: David McKay Company, Inc.

Bloom, B.S. (1986). Benjamin Bloom replies. Educational Leadership, 42 (8), 39.

Caravan, D. (1979) The influence of high and low level questions on critical thinking ability and the retention of gains in critical thinking ability effected by question types. Unpublished Master's Thesis, Memorial University of Newfoundland, St. John's.

Carpenter, H.M. (Ed.) (1963) Skill development in social studies: 33rd yearbook. Washington: National Council for the Social Studies.

Chauson, H.M. (1965) Evaluation of critical thinking in the social studies. In H.D. Berg (Ed.), Evaluation in social studies: 35th yearbook. Washington: National Council for the Social Studies.

Coffman, W.E. (1971) Essay examinations. In Robert L. Thorndike (ed.), Educational Measurement (2nd ed.). Washington, D.C.: American Council on Education.

Cohen, M.A. (1973). Teacher questioning behavior and pupil critical thinking ability: A study of the effects of teacher-questioning behavior on pupil critical thinking ability in three academic subjects offered in a suburban high school. (Doctoral Dissertation, New York University, 1972). Dissertation Abstracts International, 32, 4214A.

Cory, C.O. (1975) An experimental study to compare the development of critical thinking in eighth grade social studies classes. Saskatchewan Journal of Educational Research and Development, 6, 27-33.

Cox, C.B. (1963) Description and appraisal of a reflective method of teaching United States history. Indiana University School of Education Bulletin, 39, 74-113.

Crocker, R.K. and Riggs, F.F. (1979) Improving the quality of education: Challenge and opportunity. Final Report: Task Force on Education. Province of Newfoundland and Labrador.

Curtis, C.K. (1980) Developing critical thinking skills in non-academic social studies classes. Alberta Journal of Educational Research, 26 (2), 75-84.

D'Angelo, E. (1971) The teaching of critical thinking. Amsterdam: B.R. Gruner.

Davis, O.L. and Hunkins, F.P. (1966) Textbook questions: What thinking processes do they foster. Peabody Journal of Education, 43, 285-292.

de Bono, E. (1976) Teaching thinking. London: Maurice Temple Smith.

de Bono, E. (1983) The direct teaching of thinking as a skill. Phi Delta Kappan, 64, 703-709.

de Bono, E. (1984) Critical thinking is not enough. Educational Leadership, 42, 16-17.

de Bono, E. (1985) The CoRT thinking program. In J.W. Segal, S.F. Chipman and R. Glaser (Eds.), Thinking and learning skills, vol. 1: Relating instruction to research. Hillsdale, N.J.: Erlbaum.

Dressel, P.L. and Mayhew, L.B. (1954) Critical thinking in social science: A handbook for evaluation and teaching. Dubuque, LA: William C. Brown Company.

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

Ennis, R.H. (1980) A conception of rational thinking. In J.R. Coombs (Ed.), Philosophy of education., Normal, Ill.: Philosophy of Education Society, 3-30.

Ennis, R.H. (1984). Problems in testing informal logic/critical thinking/reasoning ability. Informal Logic, 6 (1), 3-9.

Ennis, R.H. (1985a) Critical thinking and the curriculum. National Forum, 65, 28-31.

Ennis, R.H. (1985b) Goals for a critical thinking/reasoning curriculum. Paper presented at Second International Conference on Critical Thinking and Educational Reform, Sonoma State University, California.

Ennis, R.H. (1985c) A logical basis for measuring critical thinking skills. Educational Leadership, 43, 44-48.

Ennis, R.H. and Millman, J. (1985) Cornell critical thinking tests, levels X and Z. Pacific Grove, CA: Midwest Publications.

Ennis, R.H., Millman, J. and Tomko, T.N. (1985) Manual for the Cornell critical thinking tests, levels X and Z. Pacific Grove, CA: Midwest Publications.

Ennis, R.H. and Weir, E. (1985a) The Ennis-Weir critical thinking essay test. Pacific Grove, CA: Midwest Publications.

Ennis, R.H. and Weir, E. (1985b). The Ennis-Weir critical thinking essay test manual. Pacific Grove, CA: Midwest Publications.

Fair, J.F. (1977) Skills in thinking. In D.G. Kaufman (Ed.), Developing decision making skills: 47th yearbook. Washington: National Council for the Social Studies.

Fair, J.F. and Shaftel, F.R. (Eds.) (1967) Effective thinking in the social studies. Washington, DC: National Council for the Social Studies.

Feely, T. (1976) Critical thinking: Toward a definition, paradigm and research agenda. Theory and Research in Social Education, 4, 1-19.

Fenton, E. (1967). The new social studies. New York: Holt, Rinehart and Winston.

Fraenkel, J.C. (1980) Helping students to think and value: Strategies for teaching the social studies. Englewood Cliffs, NJ: Prentice-Hall.

Fraser, D. and West, E. (1961) Social studies in the secondary school. New York: Ronald Press.

Glaser, E.M. (1941) An experiment in the development of critical thinking. New York: Bureau of Publications, Teachers College, Columbia University.

Glaser, E.M. (1985) Critical thinking: Educating for responsible citizenship in a democracy. National Forum, LXV, 24-27.

Gray, M. (1969) Research and elementary school critical reading instruction. The Reading Teacher, 22, 453-459.

Guyton, E. (1984) An analysis of the cognitive antecedents of political variables. A paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA) ERIC Document ED 245 951.

Heller, T.H. (1980) Cognitive preference testing: Process evaluation of inquiry strategies. Paper presented at the Annual Meeting of the International Interdisciplinary UAP-USC Conference on Piagetian Theory and the Helping Professions (10th, Los Angeles, CA) ERIC Document ED 184 107.

Henderson, K.B. (1958) Teaching of critical thinking. Phi Delta Kappan, 39, 280-282.

Henderson, K.B. (1972) The teaching of critical thinking. Educational Forum, 37, 45-52.

Hitchcock, D. (1983) Critical thinking: A guide to evaluating information. Toronto: Methuen.

Hodgetts, A.B. (1968) What Culture? What Heritage? Toronto: OISE Press.

Hundgins, B. (1977) Learning and thinking. Itasca, IL: Peacock Publishers.

Hunkins, F.P. (1970) Analysis and evaluation questions; their effects upon critical thinking. Educational Leadership, Research Supplement, III (April), 697-704.

Hunkins, F.P. and Shapiro, P. (1967) Teaching critical thinking in elementary social studies. Education, 88, 68-72.

Hyram, G.A. (1957) Experiment in developing critical thinking. The Journal of Experimental Education, 26, 125-132.

Kurfmann, D. (Ed.) (1977) Developing decision making skills: 47th yearbook. Washington, D.C.: National Council for the Social Studies.

Lehmann, I.J. (1963) Changes in critical thinking, attitudes, and values from freshman to senior years. Journal of Educational Psychology, 54, 305-315.

Maxim, G.W. (1977) Methods of teaching social studies in elementary school children. Toronto: Charles E. Merrill Publishing Company.

McPeck, J. (1981) Critical thinking and education. Oxford: Martin Robertson.

McPeck, J. (1984) Stalking beasts, but swatting flies: The teaching of critical thinking. Canadian Journal of Education, 9, 28-44.

Milgram, S. (1963) Behavioral study of obedience. Journal of Abnormal and Social Psychology, 67, 371-378.

Miller, J. and Weston, G.L. (1949) Slow learners improve in critical thinking. Social Education, 13, 315-316.

Morse, H.J. and McCune, G.H. (1940) Selected items for the testing of study skills and critical thinking. Bulletin 15. Washington: National Council for the Social Studies.

Newman, F.M. and Oliver, D.W. (1970) Clarifying public controversy. Boston: Little, Brown, and Co.

Norris, S.P. (1985a) The choice of standard conditions in defining critical thinking competence. Educational Theory, 35, 97-107.

Norris, S.P. (1985b) Synthesis of research on critical thinking. Educational Leadership, 42(8), 40-45.

Norris, S.P. (1986) Evaluating critical thinking ability. History and Social Science Teacher, 2, 135-146.

Norris, S.P. and King, R. (1983) Test on appraising observations. St. John's, Newfoundland: Institute for Educational Research and Development, Memorial University of Newfoundland.

Norris, S.P. and King, R. (1984) Observational Ability: Determining and extending its presence. Informal Logic, 6(3), 3-9.

Norris, S.P. and King, R. (1985). Test on appraising observations Manual. St. John's, Newfoundland: Institute for Educational Research and Development, Memorial University of Newfoundland.

Oliver, D. and Shaver, J. (1968). The effect of Student Characteristic-Teaching method interactions on learning to think critically. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Oliver, D. and Shaver, J. (1974) Teaching public issues in the high school. Utah: Utah State University Press.

Parsons, T.W. and Shaftel, F.R. (1967) Thinking and inquiry: Some critical issues. In J. Fair and F.R. Shaftel (Eds.), Effective thinking in the social studies. Washington, D.C.: National Council for the Social Studies.

Paul, R. (1982) Teaching critical thinking in the strong sense. Informal Logic, 4(2), 2-7.

Paul, R. (1984) Critical thinking: Fundamental to education in a free society. Educational Leadership, 42, 4-14.

Paul, R. (1985a) Bloom's taxonomy and critical thinking instruction. Educational Leadership, 42(8), 36-39.

Paul, R. (1985b) The critical thinking movement: A historical perspective. National Forum, 65, 2-4, 32.

Paul, R. (1985c) Critical thinking research: A response to Stephen Norris. Educational Leadership, 42 (8), 46.

Pitts-Scarangello, A.C. (1972) A study in developing critical thinking through a text in elementary school social studies. (Doctoral Dissertation, University of Pennsylvania, 1971). Dissertation Abstracts International, 32, 6296A-6297A.

Ponder, D. and Davis, O.L. (1982) Social studies education. In H.E. Mitzel (Ed.), Encyclopedia of educational research. 5th ed. New York: Free Press.

Rodnunsky, S. (1968) Changes in critical thinking scores affected by a Fenton inquiry sequence. Unpublished Master's Thesis, University of Calgary, Alberta.

Ross, J.D. and Ross, C.M. (1976a) Ross test of higher cognitive processes. Novato, California: Academic Therapy Publications.

Ross, J.D. and Ross, C.M. (1976b) Ross test of higher cognitive processes manual. Novato, California: Academic Therapy Publications.

Rothstein, A. (1961) An experiment in developing critical thinking through the teaching of American history. (Doctoral Dissertation, New York University, 1960). Dissertation Abstracts International, 21, 1141A.

Royal commission on education and youth. (1967-1968) 2 vols. Province of Newfoundland and Labrador.

Sanders, N.M. (1966) Classroom questions: What kinds. New York: Harper and Row.

Scheffler, I. (1973) Reason and teaching. New York: Bobbs-Merrill.

Scriven, M. (1985) Critical for survival. National Forum, 65, 9-13.

Shaver, J.P. (1962) Educational research and instruction for critical thinking. Social Education, 26, 13-16.

Shaver, J.P. and Oliver, D.W. (1968) The effect of student characteristic teaching method interactions on learning to think critically. Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL.)

Shipman, V. (1983a) New Jersey test of reasoning skills. Upper Montclair, NJ: Institute for the Advancement for Philosophy for Children.

Shipman, V. (1983b) New Jersey Test of reasoning skills flyer. Upper Montclair, NJ: Institute for the Advancement for Philosophy for Children.

Siegel, H. (1980) Critical thinking as an educational ideal. Educational Forum, 45, 7-23.

Sternberg, R.J. (1983) Criteria for intellectual skills training. Educational Researcher, 12 (2), 6-12, 26.

Sternberg, R.J. (1984) How can we teach intelligence? Educational Leadership, 42 (1), 38-48.

Taba, H. (1967) Implementing thinking as an objective in social studies. In J. Fair and F.R. Shaftel (Eds.), Effective thinking in the social studies. Washington: National Council for the Social Studies.

Toffler, A. (1970) Future shock. New York: Random House.

The master guide for social studies, K-XII in Newfoundland and Labrador. (1979) Province of Newfoundland and Labrador.

Unks, G. (1985) Critical thinking in the social studies classroom. Do we teach it. Social Education, 49, 240, 244-246.

Walsh, D. and Paul, R. (1986) The goal of critical thinking: from educational ideal to educational reality. Washington: American Federation of Teachers.

Warren, P.J. (1978) Public attitudes toward elementary and secondary education in Newfoundland and Labrador. St. John's: Committee on Publications, Memorial University of Newfoundland.

Warren, P.J. (1983) Public attitudes towards education in Newfoundland. St. John's: Committee on Publications, Memorial University of Newfoundland.

Watson, G. and Glaser, E.M. (1980a) Watson-Glaser critical thinking appraisal, forms a and b. Cleveland, Ohio: The Psychological Corporation.

Watson, G. and Glaser, E.M. (1980b) Watson-Glaser critical thinking appraisal manual. Cleveland, OH: The Psychological Corporation.

Weddle, P. (1986) Critical thinking in California: The department of education testing program in social studies. History and Social Science Teacher, 21, 147-157.

Westmeister, W.H. (1957) An introduction to critical thinking - A beginner's text in logic. Lincoln, NE: Johnsen Publishing Co.

Whitehead, G. (1978) Enquiry learning in social studies. Hawthorn, Victoria, Australia: Australian Council for Educational Research Limited. ERIC Document ED 164 442.

Wilen, W.W. (1985) Questioning, thinking and effective citizenship. Social Science Record, 22, 4-6.

Woods, J. and Walton, D. (1974) Informal logic and critical thinking. Education, 95, 84-86.

Wright, D.P. (1976) Interactions between instructional methods and individual aptitudes in the teaching of critical thinking in social studies. (Doctoral Dissertation, University of California, Berkeley, 1975). Dissertation Abstracts International, 37, 120A.

Wright, D.P. (1977) Instruction in critical thinking: A three part investigation. Paper presented at the Annual Meeting of the American Educational Research Association, New York. ERIC Document ED 138 518.

Wright, I. and La Bar, C. (1986) Critical thinking and social studies. In Ian Wright and Carol LeBar (Eds.), Critical Thinking and Social Studies. Toronto: Grolier Press.

Yeazell, M.I. (1981) A report on the first year of the Upshur County, West Virginia, Philosophy for children project. Thinking, 3, 12-14.

APPENDICES

APPENDIX A

(1) Letter of Introduction to Students

(2) Questionnaire for First Year Students

(3) Questionnaire for Senior Students

TO THE STUDENT

The development of critical thinking skills has long been acknowledged as a primary objective of education. Little research, however, has been conducted to ascertain the level of critical thinking abilities of students, especially those in university attendance.

At present, I am a graduate student in the Department of Curriculum and Instruction and am doing my thesis in the area of critical thinking. Your co-operation in completing the attached essay test would be greatly appreciated. Without your help, this study cannot be undertaken. Please endeavour to do your best. Thank you for your support.

Harvey Rice
M.Ed. Student
Department of Curriculum and Instruction
Memorial University of Newfoundland

General Information Questionnaire for Junior Students

Please place your respnse to each item in the space provided at the right.

1. Please indicate your sex. _____

   (1) Male

   (2) Female

2. Please indicate your age. _____

   (1) Less than 21 years of age

   (2) 21 years old or older

3. List the total number of university credits you will have at the end of this semester. _____

4. Please indicate the population of your home town. _____

   (1) 1000 or greater

   (2) Less than 1000

General Information Questionnaire for Senior Students

Please place the number corresponding to your response to each item in the response block at the right.

1. Please indicate your sex. ☐

   (1) male
   (2) female

2. List the total number of university credits you will have as of May, 1986. ☐

   (1) less than 30
   (2) 30-40
   (3) 41-50
   (4) 51-60
   (5) more than 60

3. If you have had any teaching experience, please specify length of such experience. ☐

   (1) less than one year
   (2) 1-2 years
   (3) 3-4 years
   (4) 5-10 years
   (5) more than 10 years

4. Have you taken a course(s) in Philosophy? ☐

   (1) Yes
   (2) No

5. If answer to Question 4 is yes, please specify how many. ☐

   (1) one
   (2) two
   (3) three
   (4) four to six
   (5) more than six

6. Have you taken a course specifically in critical thinking? ☐

   (1) yes
   (2) no

7. If answer to Question 6 is yes, please specify how many. ☐

   (1) one
   (2) two
   (3) three
   (4) more than three

8. If you are on the B.PE./B.A./B.Sc. and/or B.Ed. (secondary) programme, specify your major and minor areas of academic concentration.

   (1) major ..... _____
   (2) minor ..... _____

9. If you are on the B.A. (Ed.) (primary and elementary programme) specify your academic area of study.

   _____

10. List degree(s) you will hold as of May, 1986.

    _____
    _____
    _____

APPENDIX B

(1) Test Directions for the Ennis-Weir Critical Thinking Essay Test

(2) The Ennis-Weir Critical Thinking Essay Test

(3) Criteria and Scoring Sheet for the Ennis-Weir

THE ENNIS-WEIR CRITICAL THINKING ESSAY TEST
AN INSTRUMENT FOR TESTING AND TEACHING

DIRECTIONS

Read the letter to the editor of the Moorburg newspaper. Consider it paragraph by paragraph and as a total argument. Then write a letter to the editor in response to this one. For each paragraph in the letter you are about to read, write a paragraph in reply telling whether you believe the thinking good or bad. Also write a closing paragraph about the total argument. **Defend your judgments with reasons.**

Your answer should have nine numbered paragraphs. Numbers one through eight should give your reactions to paragraphs one through eight in the letter. Your paragraph number nine should give your overall evaluation of the letter considered as one total argument. Each paragraph, including the last, should contain your reason(s).

Spend about 10 minutes reading the letter and thinking about it. Then write for not more than 30 minutes (about three minutes for each of your short paragraphs). The maximum total time for the test is 40 minutes.

Do not forget to give your reasons in each paragraph. Please write clearly.

You are a local citizen, and this topic concerns you.

Remember, write **nine numbered paragraphs** and give reasons.

THE MOORBURG LETTER

230 Sycamore Street
Moorburg
April 10

Dear Editor:

Overnight parking on all streets in Moorburg should be eliminated. To achieve this goal, parking should be prohibited from 2 a.m. to 6 a.m. There are a number of reasons why any intelligent citizen should agree.

1. For one thing, to park overnight is to have a garage in the streets. Now it is illegal for anyone to have a garage in the city streets. Clearly, then, it should be against the law to park overnight in the streets.

2. Three important streets, Lincoln Avenue, Marquand Avenue, and West Main Street, are very narrow. With cars parked on the streets, there really isn't room for the heavy traffic that passes over them in the afternoon rush hour. When driving home in the afternoon after work, it takes me thirty-five minutes to make a trip that takes ten minutes during the uncrowded time. If there were no cars parked on the side of these streets, they could handle considerably more traffic.

3. Traffic on some streets is also bad in the morning when factory workers are on their way to the 6 a.m. shift. If there were no cars parked on these streets between 2 a.m. and 6 a.m., then there would be more room for this traffic.

4. Furthermore, there can be no doubt that, in general, overnight parking on the streets is undesirable. It is definitely bad and should be opposed.

5. If parking is prohibited from 2 a.m. to 6 a.m., then accidents between parked and moving vehicles will be nearly eliminated during this period. All intelligent citizens would regard the near elimination of accidents in any period as highly desirable. So, we should be in favor of prohibiting parking from 2 a.m. to 6 a.m.

6. Last month, the Chief of Police, Burgess Jones, ran an experiment which proves that parking should be prohibited from 2 a.m. to 6 a.m. On one of our busiest streets, Marquand Avenue, he placed experimental signs for one day. The signs prohibited parking from 2 a.m. to 6 a.m. During the four-hour period, there was **not one accident** on Marquand. Everyone knows, of course, that there have been over four hundred accidents on Marquand during the past year.

7. The opponents of my suggestions have said that conditions are safe enough now. These people don't know what "safe" really means. **Conditions are not safe if there's even the slightest possible chance for an accident.** That's what "safe" means. So, conditions are not safe the way they are now.

8. Finally, let me point out that the Director of the National Traffic Safety Council, Kenneth O. Taylor, has strongly recommended that overnight street parking be prevented on busy streets in cities the size of Moorburg. The National Association of Police Chiefs has made the same recommendation. Both suggest that prohibiting parking from 2 a.m. to 6 a.m. is the best way to prevent overnight parking.

I invite those who disagree, as well as those who agree with me, to react to my letter through the editor of this paper. Let's get this issue out in the open.

Sincerely,

Robert R. Raywift

Student's Name _____ Total Score _____ Graded By _____

CRITERIA AND SCORING SHEET FOR THE ENNIS-WEIR

Robert H. Ennis and Eric Weir

See manual for interpretation and qualification of these criteria.

	Credit Given (maximum is 3 points per line except #9)
1. Recognition of misuse of analogy, and/or recognition of shift in meaning, and/or claim that incorrect definition has been stipulated.	
2. Recognition of irrelevance.	
3. Recognition that Paragraph Three is OK. (Neglecting the busy-streets limitation is not penalized here.)[A]	
4. Recognition of circularity, and/or recognition that no reason is offered. (Subtract one point from credit for interpreting "undesirable" as "not desired.")	
5. Recognition that there may be other ways of preventing accidents, and/or recognition that other things might be more desirable, and/or recognition that there probably isn't much traffic at that time, and/or recognition that other types of accidents are unaffected, and/or recognition that no evidence has been given that such accidents occur. (Other possibilities)	
6. Recognition of lack of controls, and/or inadequate sampling, and/or "only one case," and/or "post hoc fallacy." (Other possible explanation)	
7. Recognition of winning argument by definition, and/or recognition that a word has been made useless for empirical assertion, and/or claim that an incorrect definition has been asserted.	
8. Recognition that Paragraph Eight is OK. (Neglecting the busy-streets limitation is not penalized here.)[A C]	
9. One point for just condemning the overall argument; another point for reviewing or summarizing the responses to the other paragraphs in some reasonable way; two points for recognizing (anywhere) the error of concluding about all streets on the basis of reasons that relate only to busy streets;[A] and one point for noting (anywhere) that Raywift has attempted to push people around with his emotive language. Total possible: 5 points.	

A score of -1, 0, 1, 2, or 3 will be given for each of the first eight numbered paragraphs:[B]

-1 judges incorrectly (good or bad)[C]
-1 shows bad judgment in justifying
0 makes no response[D]
+1 judges correctly (good or bad), but does not justify[C]
+2 justifies semi-adequately
+3 justifies adequately

For Paragraph Nine, the range is -1 to +5.

[A] Do not penalize for failure to note busy-streets limitation in Paragraphs Three or Eight. If it is not noted at least somewhere, do not give the allotted 2 points in Paragraph Nine. If the limitation is noted in Paragraphs Three or Eight, credit should be granted at Paragraph Nine.

[B] These criteria are guidelines. The grader should use judgment in awarding points, subtracting for unspecified errors and adding for unspecified insights.

[C] Sometimes, something judged one way here will be judged another way by the test taker, and so well defended that a positive score (sometimes even +3) is warranted. The grader must use judgment. For example, a good argument could be mounted against Paragraph Eight.

[D] If the examinee makes a response, but the argument of the paragraph is not judged either good or bad and no reasons are given, count it as "no response."

APPENDIX C

Robert H. Ennis' Goals for a Critical-Thinking/Reasoning Curriculum

APPENDIX

GOALS FOR A CRITICAL-THINKING/REASONING CURRICULUM

Robert H. Ennis
Illinois Critical Thinking Project
University of Illinois, U.C.
1310 South Sixth Street
Champaign, Il 61820

WORKING DEFINITION: Critical thinking is reasonable reflective thinking that is focused on deciding what to believe or do.

Critical thinking so defined involves both dispositions and abilities:

A. DISPOSITIONS:

1. Seek a clear statement of the thesis or question
2. Seek reasons
3. Try to be well-informed
4. Use credible sources and mention them
5. Take into account the total situation
6. Try to remain relevant to the main point
7. Keep in mind the original and/or basic concern
8. Look for alternatives
9. Be openminded
   a. Consider seriously other points of view than one's own ("dialogical thinking")
   b. Reason from premises with which one disagrees--without letting the disagreement interfere with one's reasoning ("suppositional thinking")
   c. Withhold judgment when the evidence and reasons are insufficient
10. Take a position (and change a position) when the evidence and reasons are sufficient to do so
11. Seek as much precision as the subject permits

12. Deal in an orderly manner with the parts of a complex whole

13. Be sensitive to the feelings, level of knowledge, and degree of sophistication of others

B. ABILITIES: (Classified under these categories: Elementary Clarification, Basic Support, Inference, Advanced Clarification, and Strategy and Tactics)

Elementary Clarification

1. Focusing on a question

   a. Identifying or formulating a question

   b. Identifying or formulating criteria for judging possible answers

   c. Keeping the situation in mind

2. Analyzing Arguments

   a. Identifying conclusions

   b. Identifying stated reasons

   c. Identifying unstated reasons

   d. Seeing similarities and differences

   e. Identifying and handling irrelevance

   f. Seeing the structure of an argument

   g. Summarizing

3. Asking and answering questions of clarification and/or challenge, for example:

   a. Why?

   b. What is your main point?

   c. What do you mean by "_____"?

   d. What would be an example?

e. What would not be an example (though close to being one)?

f. How does that apply to this case (describe case, which might well appear to be a counterexample)?

g. What difference does it make?

h. What are the facts?

i. Is this what you are saying: _____?

j. Would you say some more about that?

Basic Support

4. Judging the credibility of a source; criteria:

   a. Expertise

   b. Lack of conflict of interest

   c. Agreement among sources

   d. Reputation

   e. Use of established procedures

   f. Known risk to reputation

   g. Ability to give reasons

   h. Careful habits

5. Observing and judging observation reports; criteria:

   a. Minimal inferring involved

   b. Short time interval between observation and report

   c. Report by observer, rather than someone else (i.e., not hearsay)

   d. Records are generally desirable. If report is based on a record, it is generally best that:

      1) The record was close in time to the observation

2) The record was made by the observer

3) The record was made by the reporter

4) The statement was believed by the reporter, either because of a prior belief in its correctness or because of a belief that the observer was habitually correct

e. Corroboration

f. Possibility of corroboration

g. Conditions of good access

h. Competent employment of technology, if technology is useful

i. Satisfaction by observer (and reporter, if a different person) of credibility criteria (#4 above)

Inference

6. Deducing, and judging deductions

a. Class logic - Euler circles

b. Conditional logic

c. Interpretation of statements

1) Double negation

2) Necessary and sufficient conditions

3) Other logical words: "only", "if and only if", "or", "some", "unless", "not", "not both", etc.

7. Inducing, judging inductions

a. Generalizing

1) Typicality of data: limitation of coverage

2) Sampling

3) Tables and graphs

b. Inferring explanatory conclusions and hypotheses

1) Types of explanatory conclusions and hypotheses

a) Causal claims

b) Claims about the beliefs and attitudes of people

c) Interpretations of authors' intended meanings

d) Historical claims that certain things happened

e) Reported definitions

f) Claims that something is an unstated reason or unstated conclusion

2) Investigating

a) Designing experiments, including planning to control variables

b) Seeking evidence and counterevidence

c) Seeking other possible explanations

3) Criteria: Given reasonable assumptions,

a) The proposed conclusion would explain the evidence (essential)

b) The proposed conclusion is consistent with known facts (essential)

c) Competitive alternative conclusions are inconsistent with known facts (essential)

d) The proposed conclusion seems plausible (desirable)

8. Making and judging value judgments

a. Background facts

b. Consequences

c. *Prima facie* application of acceptable principles

d. Considering alternatives

e. Balancing, weighing, and deciding

Advanced Clarification

9. Defining terms, and judging definitions; three dimensions:

a. Form

1) Syndnym

2) Classification

3) Range

4) Equivalent expression

5) Operational

6) Example - nonexample

b. Definitional strategy

1) Acts

a) Report a meaning ("reported" definition)

b) Stipulate a meaning ("stipulative" definition)

c) Express a position on an issue ("positional", including "programmatic" and "persuasive" definition)

2) Identifying and handling equivocation

a) Attention to the context

b) Possible types of response:

i) "The definition is just wrong" (the simplest response)

ii) Reduction to absurdity: "According to that definition, there is an outlandish result"

iii) Considering alternative interpretations: "On this interpretation, there is this problem; on that interpretation, there is that problem

iv) Establishing that there are two meanings of key term, and a shift in meaning from one to the other

c. Content

10. Identifying assumptions

a. Unstated reasons

b. Needed assumptions: argument reconstruction

Strategy and Tactics

11. Deciding on an Action

a. Define the problem

b. Select criteria to judge possible solutions

c. Formulate alternative solutions

d. Tentatively decide what to do

e. Review, taking into account the total situation, and decide

f. Monitor the implementation

12. Interacting with Others

a. Employing and reacting to "fallacy" labels (including)

1) Circularity

2) Appeal to authority

3) Bandwagon

4) Glittering term

5) Namecalling

6) Slippery slope

7) Post hoc

8) Non sequitur

9) Ad hominem

10) Affirming the consequent

11) Denying the antecedent

12) Conversion

13) Begging the question

14) Either-or

15) Vagueness

16) Equivocation

17) Straw person

18) Appeal to tradition

19) Argument from analogy

20) Hypothetical question

21) Oversimplification

22) Irrelevance

b. Logical strategies

c. Rhetorical strategies

d. Presenting a position, oral or written (argumentation)

1) Aiming at a particular audience and keeping it in mind

2) Organizing (common type: main point, clarification, reasons, alternatives, attempt to rebut prospective challenges, summary--including repeat of main point)

APPENDIX D

Essentials for Effective Integration of Critical Thinking

Walsh, D. and Paul, R.W. (1986). The Goal of Critical Thinking from educational ideal to educational reality. American Federation of Teachers: Educational Issues Department, pp. 52-53.

APPENDIX

Essentials for Effective Integration of Critical Thinking

I. The role of the teacher

- a vision of teacher as professional
- valuing the teacher as critical thinker
- increasing professional autonomy
- increasing professional involvement and decision-making in policy
- revising our conception of accountability
- involvement in development of standards of practice in critical thinking

II. The education of the teacher

A. Preservice

- infusion of critical thinking throughout the curriculum
- instructors who model critical thinking practice
- intensive field experience involving observation of master teachers and supervised practice

B. Inservice

- training designed and based on the needs of teachers
- teachers involved in the development and planning of inservice
- access to coursework in the application of critical thinking in the disciplines
- quality long term staff development which capitalizes on teacher expertise and extends existing skills
- regular time to meet with colleagues to observe and learn from each other's successes and failures
- access to critical thinking materials (books, programs, skills training packages)

III. The structure of the organization

A. Commitment

- a commitment to an environment conducive to fostering critical thinking
- a realization of the magnitude of the commitment involved
- high expectations for improving and enhancing the thinking skills of all students
- a long range vision of moving toward critical thinking
- a valuing of the teachers and students as critical thinkers

B. Adequate time for critical thinking instruction

- teacher time for preparation and planning
- teacher time to consult with colleagues
- time to train staff
- adequate time to evaluate effectiveness

C. Adequate funding of critical thinking initiatives

- funds to thoroughly train staff (long term with follow-up)
- funds for teacher release time
- funds for staff to attend conferences
- funds for instructional materials
- funds for after-school committee work

D. Class size considerations

- teacher time to effectively encourage and develop thinking demands small class sizes
- time to learn how each child thinks demands opportunities for teachers to work with small groups and to provide individual assistance (teacher aides could provide this time)

E. Increased and improved standards

- consideration of critical thinking instruction in analysis of existing standards
- an emphasis on critical thinking goals translated into revised standards
- an examination of current effectiveness in developing student thinking

F. Test revision

- an analysis of critical thinking skills in existing testing programs
- a revision, where necessary, of test emphasis to incorporate critical thinking
- provisions for teacher to use essay tests in critical thinking assessment (time, aides)
- involvement of teachers in the development of assessment measures

G. Texts and instructional materials

- teacher involvement in text evaluation and selection
- education of publishers on need to incorporate critical thinking into their materials
- texts coordinated and matched with appropriate tests
- selection committees that choose materials on the basis of thinking demanded of students

H. Critical thinking committees

- appropriate support and funding for ongoing committee work
- adequate time for teachers to meet and consult with each other
- consideration and implementation of committee recommendations

I. Community involvement

- education of relevant groups on the goals and purposes of enhancing student thinking
- assistance for parents to foster cognitive growth and critical thinking in home-school partnerships
- liaisons with business and community groups to develop opportunities for real life application of critical thinking skills

J. Research considerations

- the development of future research programs and activities to support critical thinking instruction
- funding to further research and research-based training
- partnerships between schools and universities with teachers involved as action researchers

APPENDIX E

Ennis-Weir Critical Thinking
Essay Test Manual
Midwest Publications
P.O. Box 448
Pacific Grove

Comments on the Argument of the Letter and
Suggestions for Scoring (pp. 4-10)

COMMENTS ON THE ARGUMENT OF THE LETTER AND SUGGESTIONS FOR SCORING

In the following discussion of the argument of the letter and its scoring, the grader should interpret instructions for assigning and removing points as guides to be tempered by the grader's judgment.

PARAGRAPH ONE

The argument of this paragraph is a weak one. The analogy between parking overnight on the streets and having a garage in the streets is not very plausible. A related way of putting this criticism would be to say that an unconventional or incorrect meaning has been offered for the word "garage". Pointing out specific differences between parking in the streets and having a garage is a stronger criticism than merely claiming an inappropriate analogy or definition, but all are worth three points.

A more sophisticated criticism is that there is an equivocation or shift in meaning in the use of "garage" in the argument. In the first sentence, it means simply "a place to park". But in the law referred to, it most likely means (we may assume) a structure. Pointing out that the conclusion depends on this equivocation or shift in meaning is a sophisticated criticism, also worth three points.

In the light of these faults, the letter writer's failure to say where people would park their cars at night

if they did not park them in the streets is a comparatively unimportant defect of the argument of Paragraph one.

It is conceivable, though unlikely, that a respondent might argue effectively that there are important or relevant similarities between parking in the streets and having a garage in the streets (for example, occupying land). Because the ways in which they are similar are, presumably, not against the law, only partial credit (up to two points) should be given to someone defending this aspect of the paragraph's argument.

PARAGRAPH TWO

The defect of this paragraph is obvious: prohibiting parking on the streets at night will not relieve traffic congestion in the afternoon. But respondents should do more than just say that an argument is defective -- they should identify the specific defect.

Most respondents will deserve full credit for their responses to this paragraph, since the error is fairly obvious. Finer distinctions can be made by taking off credit for bad judgment or for introducing irrelevant material into a response that essentially contains a correct evaluation of the argument. For example, a respondent might accuse the writer of being biased since he obviously wants to get home in 10 minutes rather than 35. Merely having wants relevant to the conclusion does

not necessarily bias one's argument; only if the wants interfered with the reasoning would his argument be biased. So, not only is the charge of bias in the argument relatively unimportant, in this case it seems to be mistaken.

Respondents may be misled by the obviousness and simplicity of the defect of this paragraph. They may be motivated to attribute defects that it does not have. If they show bad judgment in this, this fact should not be overlooked. Take off credit. Criticism should point out the real or important defects of an argument; it should not strain to find unimportant problems or to create problems that are not actually in the argument.

PARAGRAPH THREE

The argument of this paragraph is strong -- for the streets mentioned. People on their way to work the 6 a.m. shift would be on the streets during hours Raywift proposes that parking be prohibited, and if there are no cars parked on the streets the flow of traffic will be eased. However, the argument does not fully support the specific proposal being argued for, which is to ban parking on all city streets from 2 a.m. to 6 a.m. The problem of crowded streets could be remedied by prohibiting parking on just those streets that are crowded with factory workers trying to make the 6 a.m. shift. If this limitation is not mentioned here, however, do not

take off points. Do so for the summary paragraph if the limitation is not mentioned anywhere.

An example of bad judgment in criticizing this paragraph is the accusation that "some" or "bad" are vague terms. They are undoubtedly vague terms, but (provided the limitation of the argument to "some streets" is noted) their vagueness does not seem to interfere with the strength of the argument. Unless an advantage of making them more precise is shown, such criticisms should count as bad judgment, and one point should be removed.

PARAGRAPH FOUR

As an argument for the proposition advanced by the writer, this paragraph accomplishes nothing. It cites as a reason for being opposed to overnight parking the (alleged) fact that overnight parking is undesirable. Holding something to be undesirable is not exactly the same thing as being opposed to it, but it is very nearly so; barring indications to the contrary, it can be assumed that anyone who finds a thing undesirable is opposed to it and vice versa. In other words, in this paragraph the writer has merely offered the proposition he is arguing for as a reason for itself. The argument is circular.

Another way of describing the defect of the argument of Paragraph Four is to say that it offers no additional reasons in support of the proposition being argued for.

Few respondents are likely to fail to notice the defects of the argument of this paragraph. They may, however, be tempted to attribute defects to it that it does not have. In particular, it may be asserted that the writer has simply stated his own personal opinion or preference. There is nothing wrong with stating one's personal opinions or preferences in an argument. It is to be expected.

Another mistaken objection is claiming that Raywift has not shown that overnight parking is not <u>desired</u> by the residents of Moorburg. There is a difference between claiming that something is <u>undesirable</u> and claiming that it is not <u>desired</u>. Raywift has not, on the face of it, made any claim about what is not <u>desired</u> by the residents; he has made a claim about what is not <u>desirable</u>. Respondents should be penalized one point for accusing Raywift of a fault in arguing he did not commit.

PARAGRAPH FIVE

The argument of this paragraph is not very strong, though it is the most complex of any in the letter. There are several points on which the argument can be criticized, and different respondents will either notice or give higher priority to different defects. As a result, graders might vary in their assessments of responses to this paragraph. This problem can be alleviated somewhat if graders are aware of, and give

appropriate consideration to, any challenged defect of the argument.

There are three general types of defect in the argument. The first derives from the fact that the type of accidents that would be eliminated if the recommendation were adopted is a very special and restricted type--accidents between vehicles parked on the streets overnight (or more narrowly, between the hours of 2 a.m. and 6 a.m.) and moving vehicles. Put another way, there are other, more significant types of accidents. For example, there are accidents between moving vehicles, and between moving vehicles and vehicles parked on the streets at other hours.

Another way of putting this defect is that the number of moving vehicles on the streets during the hours when parking is to be prohibited is not likely to be very great. Therefore, the number of accidents eliminated, even of this special, restricted type, is not likely to be very great. A criticism closely related to this one is that no evidence is given in Paragraph Five (or anywhere else in the letter) that accidents of the kind that would be eliminated actually do occur (or that they occur in significant numbers).

Nothing this first type of defect is a full-credit criticism.

A second type of defect has to do with the specifically evaluative aspects of the argument. It is

quite possible, for example, that other things might be judged more important than eliminating accidents of the type referred to in the argument. The inconvenience and economic costs to residents and others resulting from being unable to leave their cars on the street overnight is a consideration that is neglected by the argument. Similarly, many persons might judge the class of accidents that would be eliminated if the recommendation were adopted to be relatively unimportant compared to those that would remain unaffected.

Finally, there is a defect that arises from the form of the argument. There are several ways in which this type of defect might be described. In ordinary, non-technical language, the defect is that the argument describes only one possible way of eliminating accidents of the type it claims would be eliminated. Since there may be other ways of eliminating such accidents, it is not incumbent even on those who agree that it is important to eliminate them to support adoption of this specific remedy. In logical terms, the letter writer has shown the recommendation to be a probable sufficient condition for eliminating one kind of accident, but not a necessary condition. Another way of putting this is to say that (roughly speaking) the argument commits the fallacy of affirming the consequent: from the fact that accidents will be prevented if parking is eliminated, and the desirability of eliminating accidents, it does not

necessarily follow that parking should be prohibited; there may be other ways of eliminating accidents. Noticing this defect, however it is described, indicates greater sophistication than noticing the other types of defect.

Each of these three types of defect is an "other-possibilities" defect. That is, in each case there is another reasonable possibility. So Paragraph Five is an other-possibility paragraph.

Although the argument of this paragraph is properly regarded as not very strong, it is possible for a response that judges the argument to be good to deserve full credit. This will generally be when the respondent qualifies the judgment that the argument is good by taking note of any of the features we have just identified as defects. For example, a respondent might say, "The argument gives a reason for prohibiting parking, but probably few accidents would actually be prevented, and it isn't shown that we should <u>necessarily</u> be in favor of the proposal." It is important to preserve the possibility of judging the argument to be good because many good arguments will have the same structure as this weak one. For example, if there was general agreement that some outcome should be avoided, and if other ways of avoiding it were either unavailable or undesirable, then an argument of this form would provide a reason for

supporting or adopting the means referred to in the antecedent of the argument.

PARAGRAPH SIX

Paragraph Six can be regarded as an other-possible-explanation paragraph. The most important defect of the argument is that the results of the one-day experiment do not adequately support the causal claims implicit in it. The argument implies that the lack of accidents in the four-hour period the day of the experiment was due to the installation of no-parking signs. It also suggests that, if parking were prohibited on other streets, accidents would again be prevented. There are, however, a variety of plausible alternative explanations for the lack of accidents. The existence of these other possible explanations also undermines the generalization of the experimental results. Since the argument does not present sufficient information to rule out these other explanations, it does not provide adequate support for the claims being argued for.

For example, it might be claimed that the lack of accidents during the period parking was prohibited could easily have been simply a chance occurrence and not really due to the parking ban itself. The fact that there were no accidents during such a brief period would not be a surprising occurrence. The inference that installing the no-parking signs was effective in eliminating accidents

would have been stronger if observation had been extended over a significantly longer period.

Another possibility is that the day on which the experiment was conducted could easily have been atypical in some way. We are not told that the experiment was conducted on a normal work day. For all we know, it was conducted on a weekend or holiday, or perhaps there was a bad snowstorm that day. If any of these propositions were rue, the reduced volume of traffic the day of the experiment would be a plausible alternate explanation of the results. Still another potential explanation arises from the possibility that there were no accidents on the streets where signs were not installed. If that were the case, not only would it be reasonable to suspect that something other than the no-parking signs was responsible for the lack of accidents on Marquand, we would also have reason to doubt that extending the parking prohibition to other streets would eliminate any more accidents.

A final alternative arises from the fact that we are not told how many of the more than 400 accidents occurring on Marquand involved parked vehicles. If few of them did, we would have to seek an explanation for the lack of accidents in some other factor, since the parking prohibition would probably not be responsible for eliminating accidents that did not involve parked vehicles in the first place. For the same reason, it would be

unreasonable to expect that prohibiting parking on other streets would prevent accidents.

It is important to note that these are all explanations that are reasonable to propose. There are many other explanations that could conceivably be offered, but they would not necessarily constitute valid criticisms of the experiment or the conclusions drawn from it. For example, it is conceivable to suggest that alien beings hovering in a nearby spaceship intervened in some way to prevent the accidents. But clearly we would not take such a suggestion seriously. Only when the alternative explanation is a <u>reasonable</u> one to propose does it constitute a significant criticism of the experiment.

Some respondents may fault the argument for its use of the expression "everyone knows, of course," on the ground that this is an attempt to exercise unwarranted influence on the reader. This is a weak criticism and should be faulted for exhibiting bad judgment. If there were this many accidents in one year, it would not be unreasonable to assume that most people knew about it. And in any case, it is an easily checked factual claim. It does not appear that Raywift has attempted to gain unwarranted assent in asserting that "everyone knows."

An adequate response to the paragraph would judge the argument to be weak and would indicate in some way that the reason for this is the tenuousness of the implied causal claims. A criticism that is justified and is at

least fairly specific receives full credit. For example, even if the response says simply, "The experiment does not prove that prohibiting parking caused the lack of accidents," it should be given full credit. As the criteria indicate, the same fundamental criticism may take a variety of forms and be expressed in many different ways.

A problem that might not be noticed by unsophisticated respondents is that the claim being argued for, a value claim (i.e., that parking should be prohibited) has not been adequately supported, the problems about causation aside. When this problem is noticed, the appropriate criticism would be similar to those applied to Paragraph Five (e.g., that the inconvenience of the parking ban would be too great or that there are other ways of preventing accidents). Respondents who make such a criticism should not be penalized for not mentioning the problems about causation. They should receive full credit.

PARAGRAPH SEVEN

The defect in Paragraph Seven can be put several different ways. It might be claimed that the definition that is stipulated is simply incorrect, that this is not what "safe" really means. Another way of challenging the defect is that the proposed definition actually renders the word useless, since not every chance of an accident

can be eliminated. Not only would present conditions be unsafe if the proposed definition were accepted, they would remain so even if Raywift's proposal were adopted. Thus his definition, though framed to suit his purpose, is actually self-defeating. A third way is to note that the meaning of the word "safe" has been shifted in mid-argument, making the argument a case of equivocation.

An adequate response will at least judge the argument weak. If the justification is that the definition is incorrect, the response should be given three points. This is a reasonable criticism, though not as incisive as pointing out that the word has been rendered useless, or that equivocation has occurred. Either of the latter criticisms are also worth three points.

An example of poor judgment is justifying a correct evaluation of the argument would be a claim that Raywift's definition is vague or that it is unclear what he means by "safe." His definition is very clear; it just cannot actually be satisfied. This criticism should cost the test taker one point.

If the respondent correctly judges the argument, but justifies the judgment only by claiming that Raywift has "slandered" his opponents in accusing them of "not knowing what safe really means," give credit only for the correct judgment (one point). While Raywift's claim about his opponents' knowledge of the meaning of the word may be

intended to have rhetorical effect, this is a trivial criticism compared with those mentioned above.

PARAGRAPH EIGHT

The argument of this paragraph is one of the better ones in the letter. The authorities cited could reasonably be expected to be knowledgeable about the subject being discussed. Their recommendation is directly relevant to Raywift's proposal. Further, there seems to be no good reason to doubt their expertise or to doubt that they actually made the recommendation he claims they made. There is, however, a crucial qualification that weakens the support provided for Raywift's proposal: the authorities' recommendation applies only to busy streets. Again, do not take off credit if this limitation is not noted here.

Generally, an adequate response would judge the argument to be reasonably good. If the citing of these particular authorities is judged to be relevant and appropriate, and to lend force to the argument, the response should be given three points. This judgment should be explicit, however. If the respondent supports a positive judgment only by stating that authorities are cited, without commenting on the relevance or appropriateness of these particular authorities, he or she should be given credit for a semi-adequate justification only (two points). On the other hand, "The author appeals to two

different legitimate authorities" or "This is all right, if the authorities are qualified" would be marginally worth three points, because they indicate concern for the authorities' qualifications.

Raywift is not required to give evidence that the authorities actually said what he claims they have said, or to produce the reasons they gave for making their recommendation. Respondents who criticize him for not doing either of these things should be faulted for bad judgment in justifying. The important point is that their recommendation is relevant to his. The fact that they made it can be checked if necessary. If the reasons offered in arguments were generally judged suspect when not themselves positively justified, almost all reasons in real-life arguments would be under suspicion. Claims made in the course of an argument, if they can be easily checked, should generally be granted credibility--unless there are reasons for doubting their truth or relevance.

Some respondents, however, may wonder whether the authorities' recommendation would still appear relevant to Raywift's purpose if their reasons were known; they may point out that his argument would have been stronger if he had indicated their reasons and shown them to be relevant to his purpose. Where possible, this should be distinguished from demanding positive justification for the authorities' recommendation or evidence that they actually made it, and respondents should be given credit

for observing that the strength of the argument could be affected by knowledge of the authorities' reasons.

A respondent might judge the argument in Paragraph Eight weak on the ground that it is dangerous to infer from a loose general recommendation to this particular city without knowing that Moorburg is typical of cities this size, or that the facts about Moorburg do not disqualify it from fitting this general recommendation, because Moorburg might well be different from other cities its size. Give full credit for this sort of sensitive skepticism.

Some respondents may judge the argument weak because they take it to be advancing a trivial, tautologous claim. It may be argued that it is "obvious" that the best way to prevent overnight parking is to prohibit parking from 2 a.m. to 6 a.m.--as if the two were equivalent. They are not equivalent: prohibition of parking between 2 a.m. and 6 a.m. is offered as a means of discouraging overnight parking, and Raywift's proposal makes this clear. Take off one point when this criticism is made.

If a respondent rejects the argument as an "appeal to authority," apparently believing that any appeal to authority is fallacious, then give a minus one.

PARAGRAPH NINE

Responses in this paragraph are among the most difficult to rate in the test. What is desired is a

judgment of the overall strength of Raywift's argument that gives specific reasons but does not simply recapitulate the responses of the first eight paragraphs.

The minimum requirement of an adequate response is that the argument of the letter be judged faulty (worth one point). It is difficult to imagine that a plausible case could be made for a judgment to the contrary. To receive more than one point, however, the response should do more than just condemn the argument (by merely calling it "fallacious," for example). If, in addition, the response says that six of the eight paragraph contain faulty arguments, if it correctly identifies the two paragraph with reasonably good arguments, or if it simply summarizes the judgments made in the preceding paragraphs, give it one more point.

If, and only if, the mistake of inferring from some streets to all streets is mentioned somewhere, even if it is not mentioned here, the respondent should receive two more points.

The use of emotive language in the introductory paragraph ("any intelligent citizen") is an attempt to get people to agree by illicit means. Noting this sort of thing (somewhere) is good for one point (here),

So, five points are available from Paragraph Nine.

APPENDIX F

Correspondence

46 Cashin Avenue
St. John's
Newfoundland
Canada
A1E 3A8

September 5, 1986

Mr. John Baker
Midwest Publications
P.O. Box 448
Pacific Grove CA 93950
U.S.A.

Dear Mr. Baker:

At present, I am a graduate student in the Department of Curriculum and Instruction at Memorial University of Newfoundland working, under the supervision of Dr. Frank Cramm and Dr. Stephen P. Norris, on a thesis tentatively titled An Examination of the Critical Thinking Ability of Entering and Leaving Education Students.

I would like to request permission from your company to include a copy of the Ennis-Weir Critical Thinking Test: An Instrument for Testing and Teaching, as an Appendix to my thesis. I feel that the inclusion of the test, test manual and criteria and scoring sheet would be a valuable aid to readers of the thesis in understanding not only the instrument itself, but also the basis on which students' responses were evaluated.

Thank you for your consideration of this request.

Yours truly,

Harvey Rice

MIDWEST
PUBLICATIONS
THINKING SKILLS
BOX 448, PACIFIC GROVE,
CA 93950 (408) 375-2455

September 22, 1986

Harvey Rice
46 Cashin Avenue
St. John's, Newfoundland
A1E 3A8
CANADA

RE: REPRODUCTION OF "ENNIS-WEIR CRITICAL THINKING TEST"

Dear Mr. Rice:

In responding to your letter of September 5, 1986, you have our permission to quote pages from the Ennis-Wier Critical Thinking Essay Test as an Appendix to your thesis just as long as credit is given to us.

Thank you for notifying us for permission. If you should need any other assistance please let us know and we will be happy to help you.

Sincerely,

MIss Michelle L. Carlsen
Order Dept.

MLC/eht