An Exploratory Attempt to Measure Certain Comprehension Skills Necessary for the Critical Reading of Two, Topic-Related Expository Selections

Patricia F. Borowiec

The College at Brockport

Follow this and additional works at: https://digitalcommons.brockport.edu/ehd_theses

Part of the Higher Education Commons, and the Language and Literacy Education Commons

To learn more about our programs visit: http://www.brockport.edu/ehd/

Repository Citation


https://digitalcommons.brockport.edu/ehd_theses/952

This Thesis is brought to you for free and open access by the Education and Human Development at Digital Commons @Brockport. It has been accepted for inclusion in Education and Human Development Master’s Theses by an authorized administrator of Digital Commons @Brockport. For more information, please contact kmyers@brockport.edu.
AN EXPLORATORY ATTEMPT TO MEASURE CERTAIN COMPREHENSION SKILLS
NECESSARY FOR THE CRITICAL READING OF TWO, TOPIC-RELATED,
EXPOSITORY SELECTIONS

THESIS

Submitted to the Graduate Committee
of the Department of Curriculum and Instruction
Faculty of Education
State University College at Brockport
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Education

by

Patricia F. Borowiec
State University College at Brockport
Brockport, New York
December, 1978
Abstract

The purpose of this study was to describe the results of a researcher-constructed, critical reading assessment administered to 224 college students (133 first-year students and 91 third-year students) enrolled in a three-year degree program. Subjects were tested on two, longer, topic-related, expository selections.

Statistical analysis indicated that total performance score and type of conclusion reached about the selections read were significantly related to the ability to recognize overgeneralizations and contrasting views. Analysis also indicated that those subjects (a) with the higher academic averages, (b) who were humanities majors, (c) had a New York State regents diploma, and/or were college seniors scored highest on this assessment.

Finally, Gutman scaling supported the difficulty ranking for the opinion/interpretation items. It did not support the rankings for the literal comprehension and contrast items.
# Table of Contents

List of Tables ................................................. A

Chapter

I Statement of the Problem ................................. 1
  Purpose ..................................................... 1
  Questions to be Answered ................................. 1
  Need for the Study ....................................... 2
  Definitions of Terms ..................................... 4
  Limitation of the Study ................................. 5

II Review of the Literature ................................. 7
  Purpose ..................................................... 7
  Areas of Research ....................................... 7
  Critical Reading and its Subskills ...................... 7
  The Need for Proficiency in Critical Reading .......... 12
  Previous Attempts to Measure Critical Reading Proficiency ........................................... 14
  Summary .................................................... 22

III Design of the Study .................................... 24
  Purpose ..................................................... 24
  Subjects .................................................... 24
  Instrument .................................................. 24
  Description of Test Items ............................... 25
  Statistical Analysis ..................................... 28
  Summary ..................................................... 28

IV Statistical Analysis .................................. 29
  Purpose ..................................................... 29
  Findings .................................................... 29
  Comparison of Subjects by Type of Conclusion Reached .................................................. 33
  Difficulty Rankings of Literal, Opinion/Interpretation and Contrast Items .................... 41
  Academic Performance and Assessment Performance ......................................................... 43
Table of Contents (continued)

| Type of New York State Diploma and Assessment Performance | 43 |
| College Major and Assessment Performance | 43 |
| Sex and Assessment Performance | 52 |
| Grade Level and Assessment Performance | 53 |
| Summary | 54 |

V Conclusions and Implications ................. 56

Purpose ........................................ 56
Conclusions .................................... 56
Implications for Research ..................... 60
Implications for Classroom Practice .......... 61

References .......................................... 63

Appendices

A. Extended Lists of Critical Reading Skills .. 69
B. Researcher-constructed Assessment .......... 71
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Pearson's Correlation Between Subskill Category Performance and Total Performance Score: (a) Excluding Type of Conclusions and (b) Including Type I Conclusion</td>
<td>32</td>
</tr>
<tr>
<td>2.</td>
<td>Analysis of Variance of Means on Literal Questions by Conclusion Type Groups</td>
<td>34</td>
</tr>
<tr>
<td>3.</td>
<td>Analysis of Variance of Means for Identification of Factual Statements by Conclusion Type Groups</td>
<td>35</td>
</tr>
<tr>
<td>4.</td>
<td>Analysis of Variance of Means for Identification of Opinion/Interpretation Statements by Conclusion Type Groups</td>
<td>36</td>
</tr>
<tr>
<td>5.</td>
<td>Analysis of Variance of Means for Recognition of Overgeneralizations by Conclusion Type Groups</td>
<td>37</td>
</tr>
<tr>
<td>6.</td>
<td>Analysis of Variance of Means for Recognition of Comparisons by Conclusion Type Groups</td>
<td>38</td>
</tr>
<tr>
<td>7.</td>
<td>Analysis of Variance of Means for Recognition of Contrasts by Conclusion Type Groups</td>
<td>39</td>
</tr>
<tr>
<td>8.</td>
<td>Analysis of Variance of Mean Total Performance by Conclusion Type Group</td>
<td>40</td>
</tr>
<tr>
<td>9.</td>
<td>Academic Performance and Assessment Performance Means</td>
<td>45</td>
</tr>
<tr>
<td>10.</td>
<td>Academic Performance and Type of Conclusion Reached</td>
<td>47</td>
</tr>
<tr>
<td>11.</td>
<td>Type of Major and Assessment Performance Means</td>
<td>50</td>
</tr>
</tbody>
</table>
Chapter I

Statement of the Problem

Purpose

The primary purpose of the present study is to determine the content validity of a researcher-designed, critical reading assessment. This assessment is an attempt to measure six sub-skills involved in the critical reading of two topic-related, expository selections. The secondary purpose of the study is to describe the performance of subjects to whom the assessment was administered.

Questions to be Answered

1. Does the assessment produce a normal range of variability in the total performance scores?

2. Is there a relationship between the type of conclusion drawn by subjects and their performance scores for other sub-skill categories?

3. Is there a relationship between performance on more difficult items and performance on items of less difficulty for the skill categories of literal comprehension, recognition of contrasts, and opinion/interpretation recognition?

4. Is there a relationship between a subject's total performance score and other factors, such as scholastic performance,
grade level and sex?

5. Is there a relationship between a subject's subskill category performance scores and other factors such as scholastic performance, grade level and sex?

Need for the Study

There is no question among reading specialists as to the importance of reading critically. Definitions of critical reading, however, vary greatly (Huus, 1971; Marksheffel, 1966; Russell, 1959; Williams, 1969). Accordingly, the lists of subskills necessary for critical reading vary just as much, especially if a distinction is not made between the critical reading of expository material and the critical reading of fiction.

There have been several recent attempts to measure critical reading. Instruments used to measure this skill have been either investigator-designed or borrowed from a previous researcher in this area. The subskills of critical reading included in these instruments have varied just as the previously mentioned lists vary. In describing the results of their studies, these researchers have reported the performances of subjects according to categories such as sex, age and intelligence, concluding that there is or is not a significant relationship between critical reading ability and these other factors. There appears to be
little concern for the content validity of their instruments — whether it has indeed measured the skill and/or subskills of critical reading.

Aside from the abovementioned deficiency in validation of instruments, a basic problem appears to be one of trying to measure too much with too little. The Ohio State University Critical Reading Test (1967), for example, attempts to measure twenty factors (subskills) involved in the critical reading of expository and/or fictional material. Additionally, of those twenty factors, fifteen are measured by only one or two items.

There is a multitude of subskills involved in the critical reading of fiction and an equal number involved in the critical reading of expository materials. These skills do not always overlap. The subskills involved in critically reading Konrad Lorenz's On Aggression are not the same as those necessary to critically read Randall Jarrell's The Death of the Ball Turret Gunner. Furthermore, not all fiction and expository pieces demand proficiency in all of the subskills related to each type of writing. A particular expository article may require that the reader be able to recognize opinion and overgeneralization, but not inference because the author has explicitly stated his views.

Therefore, it does not appear unreasonable for several critical reading tests, of two major types, to exist. The first type of tests would measure only those subskills related to
fictional writing. The second type would measure only those subskills necessary for the critical reading of expository writing.

The present study is an attempt to devise a test of the second type. The subskills to be considered are the abilities to: (1) comprehend literally; (2) recognize overgeneralizations; (3) recognize fact and opinion/interpretations; (4) compare; (5) contrast; and (6) reach a conclusion through evaluation of what has been read.

Definitions of Terms

content validity - the extent to which the content of the test represents a balanced and adequate sampling of the outcomes of the course or instructional program it is intended to cover.

critical reading - the subskills literal comprehension, recognition of overgeneralizations, recognition of fact and opinion/interpretation, comparison, contrasting and reaching a conclusion through evaluation of what has been read.

literal comprehension - the ability to correctly respond to questions, the answers to which are explicitly stated in the text.

overgeneralization - inflexible statements of assertion that
offer no exceptions to hard and fast rules or principles.

factual statement - a statement, the components of which may be tested objectively to determine their validity.

interpretation/opinion - value judgments made by the observer of an event or events. These may be defined separately, but are often popularly used interchangeably and have been combined on this assessment for that reason.

compare - to observe the similarity(ies) of two authors' discussions on the same topic.

contrast - to observe the difference(s) between two authors' discussions on the same topic.

conclusion - The last part of a chain of reasoning; judgment, decision, or opinion formed after investigation or thought.

Limitations of the Study

Because this is a researcher-constructed assessment, the following conditions exist:

1. it is not a standardized test;

2. it was not constructed by specialists in the area of critical reading, although the researcher did consult with professors of reading education.
Therefore, because this is a pilot study, the description of results will be confined to assessing the content validity of the instrument and describing the performance of subjects involved in the study.

Also, there is only one item attempting to measure the reader's ability to reach a critical conclusion in expository materials. However, because a subject must state the reasoning behind the conclusion drawn, it is hoped that this single item will indicate a particular subject's ability to reach a critical conclusion when reading expository materials which involve the same subskills measured by this assessment.

**Summary**

To this researcher's knowledge, there does not exist an instrument with which to measure a reader's ability to critically read two topic-related, longer, expository selections. If reading specialists are in agreement that critical reading of related expository material is a skill important for an adult to function effectively in society (Cameron, 1968; Harris, 1975), then it appears equally important that there be a means for determining whether or not this person has acquired this skill by the time formal schooling is nearing completion.

This is a small area of the total field of critical reading, and this assessment is an attempt to devise an instrument which will measure proficiency within that area.
Chapter II

Review of the Literature

Purpose

The primary purpose of the present study is to determine the content validity of a researcher-constructed, critical reading assessment. This assessment is an attempt to measure six subskills involved in the critical reading of two topic-related, longer, expository selections. The secondary purpose of the study is to describe the performance of subjects to whom the assessment was administered.

Areas of Research

Three areas of research appear relevant to this study:

1. definitions of critical reading, and the subskills considered pertinent to critical reading;

2. the expressed need for proficiency in critical reading and, by extension, the need for instruments which measure the degree of proficiency in this skill; and

3. the results of previous attempts to measure proficiency in critical reading.

Critical Reading and its Subskills

As previously mentioned, definitions of critical reading vary. An author's definition of this skill appears dependent upon whether or not he is concerned with the reading of fic-
tional materials, expository selections, or both. These concerns are in turn affected by whether or not the author is: (a) concentrating upon certain subskills involved in critical reading which he feels have been neglected; (b) proposing a strategy for teaching certain subskills at specific grade levels; or (c) attempting to measure the entire area of critical reading. In some instances, the author's definition is not explicitly stated; but rather, it is inferred by the listing of subskills thought pertinent to the act of critical reading.

The general consensus, however, is that critical reading is the act of evaluating and passing judgment upon what is read, based upon some known standards or norms. This broad definition is applicable to both expository and fictional materials. It is the amplification of this definition that appears to create the inconsistencies in defining critical reading and its subskills to which Lundsteen and Hackett (1969), Wolf, Huck & King (1967), and Robbins (1977) have referred.

Wardeberg (1967) said the critical reader must be an "artist of appreciation," having developed taste, discrimination and insight. Harris (1975) defined critical reading as involving the "comparison of two or more sources of information... considering new ideas or information in the light of one's previous knowledge and beliefs...and the ability to detect and resist the influences of undesirable propaganda." (One
wonders how Harris defines desirable propaganda.) Russell (1963) stated that the critical reader simply applies critical thinking to the act of reading. He defined critical thinking as involving questioning and reserving judgment, logical inquiry and problem-solving, and evaluating by some norm or standard. "It is," according to Russell, "a logical examination of data which avoids fallacies and judgments on an emotional basis only."

Taschow (1972) stated that the critical reader must react emotionally as well as intellectually. The Montgomery County Public Schools report (1974) defined evaluation as being both subjective and objective. Robinson (1964) defined critical reading as the "judgment of the veracity, validity, or worth of what is read, based upon sound criteria or standards developed through previous experiences."

A source of inconsistency appears to lie in how one defines norms or standards. All agree that the norms or standards are derived from previous experience. Heilman (1970) stated that "it is likely that the greatest barrier to critical reading is the reader's lack of background and experience." But he also stated that the "reader's previous experiences or attitudes interfere with the analysis and evaluation of what he reads."

In the evaluation of fictional material, the reader may judge the actions of a character as "good" or "bad" based upon his own values and/or how he thinks he or someone he knows would react in a similar situation. These values and the assessment
of how someone other than the character might react are based upon the reader's previous experiences. Furthermore, the reader may accept or reject the works of a fictional writer based upon some preferences for particular styles of writing. These preferences also are based upon previous experiences.

In this reading situation, the reaction is subjective. The reader is expected to react personally (Williams, 1959; Jacobs and Searfoss, 1977).

The previous experiences for the critical reader of expository materials, however, are of a different type. In this reading situation, personal attitudes interfere with the reader's evaluation of what is being read. The previous experiences in this case are factual knowledge about the subject being discussed which will help in evaluating the validity of what has been written. The standards or norms are objective and based upon logic.

In defining critical reading, many authors do not differentiate between the two types of writing and the personal experiences as well as individual subskills indigenous to the critical reading of each. The result of this lack of differentiation is confusion as to how critical reading is defined and exactly what subskills are involved (King, 1968; Jenkinson, 1969; and Robbins, 1977).

Accordingly, the lists of subskills vary in length and specificity. Either there is an attempt to list all possible
subskills without indicating which subskills apply to fictional material and which apply to expository material, or the list contains a few subskills which relate to one or the other type of writing without acknowledging which type, and/or that the list is a partial one.

DeBoer (1946) said the critical reader must be active, distinguish between relevant and irrelevant material, and be skeptical in order to evaluate evidence and conclusions. Smith (1963) included: an inquiring attitude; willingness to read as long as it takes to find the truth; evaluation; “deciding upon truthfulness, bias, authenticity”; and reacting personally to what the author has stated “based upon experience, facts gleaned from various sources, or possibly as a result of clear-cut reasoning.”

Cohen (1967) included four subskills: identifying setting; character development; thematic approach; and analogies. McCord (1967) asked that the reader recognize bias and look at: the qualifications of the author; the authenticity of factual material; and the financial source for the writing and publishing of the material.

Morse and McCune (1971) included: distinguishing between statements of fact and statements of motive; recognizing bias; drawing inferences; and evaluating sources of information.

Biberstine (1975) listed the subskills of: comparing and contrasting; collecting and organizing information; and decision-
The Montgomery County Schools report listed the following as critical reading/critical thinking skills: (1) word meanings; (2) location/recall; (3) translation; (4) interpretation; (5) prediction; (6) application; (7) analysis; (8) synthesis; and (9) evaluation. Roe (1978) included "the need to sort out fact from opinion, truth from half-truth, information from emotion."

Other, longer lists of critical reading subskills are included in Appendix A. The first of these lists (Raciti, 1969) includes two skills which appear contradictory: "understand need to suspend judgment until more information is known"; and "be willing to accept the ideas of the author if he is an expert." The other two lists are the most inclusive and specific. However, they, with the others, do not differentiate between which subskills apply to fictional writing and which apply to expository.

The Need for Proficiency in Critical Reading

There is no dispute among reading specialists as to the importance of being able to read critically. Proficiency in this skill is deemed necessary for the adult to function successfully within society. To function successfully is to be intelligently involved in the decisions which affect one's daily existence on the public and private levels (Cameron, 1968).

McCord saw critical reading as a vital skill without which the reader is unable to use effectively what has been read. Both Wolf and Roe stated that, with the yearly increase in the quantity of printed materials, the need for critical reading increased proportionately. Bloom saw it as the responsibility of every adult to evaluate and "solve the problems of the world. In a democracy, the need for evaluation is apparent."

Using Piaget's definition of intelligence, Cahn stated that in order for an adult to comprehend the world in which one lives and "to build or discover new structures within it," that person must have the ability to solve problems. And the ability to solve problems, according to Cahn, is dependent upon the skill of critical reading.

Other specialists have implied the need for critical reading through their suggested strategies and/or their call for the teaching of this skill at the various grade levels (Sailor, 1967; D'Angelo, 1971; Duquette, 1973; Greer, 1973; and Queen, 1973).

Until recent years, reading teachers/specialists have considered critical reading to be at the top of the reading skills ladder. The reasoning was that because the reader must be able to decode and understand literally before reading critically, then that reader must have an abundance of the
former two skills before attempting the latter. Most of the research prior to 1967 indicates that critical reading was the concern of teachers and researchers as it applied to high school students (King, 1968).

In the last ten years, the concern for teaching critical reading at the secondary level has not diminished. There has been, however, a re-evaluation of how early in the reading program some of the subskills should be taught (Wolf, 1968; Raciti, 1969; and Raschow, 1972). It is generally agreed that the teaching of critical reading can and should begin in the primary grades. Additionally, the literature shows a demand for the continuation of instruction through the college level (Lee, 1967; Stanton, 1969; Downey, 1974 and Halliden, 1977).

Previous Attempts to Measure Critical Reading Proficiency

Several attempts have been made to measure the critical reading proficiency of students grouped according to sex, intelligence, and grade level. Other studies have compared critical reading ability to general reading ability and/or various subskills involved in comprehension of reading materials. One has investigated the relationship between critical reading and critical thinking; another has studied the effects of reading rate and anxiety upon critical reading ability. Other related studies considered regional differences in this skill and the effects of authority-based statements upon critical reading.
Gall (1973) investigated the relationships between critical reading skill and sex, intelligence and general reading achievement, as well as possible differences in proficiency between students at the fourth, fifth, and sixth grade levels. She found: a significant difference in critical reading ability between grade levels; a strong relationship between critical reading and intelligence, and general reading ability and critical reading; and no significant difference between the sexes. Her instrument was a 30-item, investigator-devised assessment.

Johnson's study (1974) also looked at grade level and sex differences. Again, grade level was found to be significant and sex not. The instrument was researcher-constructed.

The Wolf, Huck and King study (1967) attempted to determine whether or not critical reading could be taught at the elementary grade levels. (A complete list of subskills included on their instrument - The Ohio State University Critical Reading Test - can be found in Appendix A. However, there are three forms of the test, and each form does not include all of the subskills.) The experimental group at each grade level was instructed in the critical reading of fictional material, while each control group worked in a basal reader program.

Grades one through six experimental groups did consistently better than the control groups on the logic section of
of the posttest. Only the grade three experimental group did significantly better than the grade equivalent control group on all three sections of the test. The performance on the logic section is interesting in light of Lundsteen and Hackett's (1969) reference to how little time was spent on reading to develop the skills involved.

As with the Gall study, the Wolf, Huck & King research indicates a high correlation between intelligence and critical reading ability. However, in comparing low IQ subjects in the experimental and control groups, this study found an increase in subskills for the experimental group which would indicate that low intelligence does not preclude increasing critical reading ability.

An earlier observation in reference to the Wolf, Huck & King study should be considered. Of the twenty subskills included on one form of the instrument, fifteen of these were measured by only one or two items. Additionally, one might question how each subskill is defined by the test item used to measure that subskill. An example of this is an item intended to measure the ability to contrast two paragraphs which discuss seashells. The subject is asked to choose which answer describes the two paragraphs. Three of the four possible answers describe one or the other paragraph erroneously. This item appears to measure the ability to recognize main idea rather than contrasting.
Heiman (1976) used the Wolf, Huck & King instrument to measure the effects of instruction in logical thinking upon fourth, fifth, and sixth graders. She concluded that the treatment had a significant effect ($p < .001$), but questioned whether the effect was related to the concepts taught, materials used, the method of instruction, or some combination of these factors. She also found that "the effect of grade and sex related to children's scores was unclear."

Robbins (1977) considered the relationships between critical reading and "selected measures of literal and interpretive reading." Her test materials included unrelated paragraphs from social science texts. Skills included as literal were: main idea; noting details; and vocabulary. Interpretation skills were: recognizing point of view; distinguishing between fact and opinion; inference; pertinence ("ability to select data which are pertinent to support the author's point of view"); competence (of the author); and sufficiency ("whether the author presents sufficient data to support his point of view"). Critical reading was defined as the reader's ability to "judge the truthfulness of the author's point of view." To measure this skill, the subject was asked to judge the degree of truthfulness of a series of statements related to a paragraph previously read.

The strongest relationship was between interpretive skills proficiency and critical reading. "The association of the
literal with the critical was controlled by the interpretive, but the relationship of the interpretive with the critical was not controlled by the literal." The following subskills were found to be significantly related to critical reading: main idea; point of view; pertinence; competence; sufficiency; and distinguishing between fact and opinion. The strongest relationship was with fact and opinion, with sufficiency being the next strongest.

Robbins also administered the paragraph meaning section of the *Stanford Achievement Test* for a general reading score. No significant relationship was found between that score and critical reading ability.

Sochor (1958) and Maney (1958) also found no significant correlation between general reading ability and critical reading proficiency. Sochor found a .23 correlation using science text materials, while Maney arrived at a .11 correlation using material from a social studies text. Maney also found a .67 correlation between intelligence scores and critical reading ability, confirming Sochor's earlier similar correlation for these two variables.

Gans (1940) found no relationship between a subject's literal comprehension and critical reading ability. Her assessment of critical reading proficiency measured the subject's ability to distinguish between relevant and irrelevant material. The subject was given a problem to solve and asked to select
among an assortment of source materials those which were relevant to the solution of the problem. She found that a majority of the intermediate grade subjects in her study were unable to distinguish between relevant and irrelevant materials.

Van Voorhes (1974) administered a critical thinking test to determine the effects of rate and anxiety upon critical reading. His slow and fast readers all had an IQ of 120 or higher and all had scored above the median on the vocabulary and comprehension sections of a standardized test. Under timed and untimed conditions, Van Voorhes found that fast readers were more critical readers. But low anxiety readers scored significantly higher than high anxiety readers, regardless of rate. Van Voorhes justified the use of a critical thinking appraisal by saying that such an appraisal "necessitates critical reading."

Downey (1976) investigated "critical thinking while reading." His instrument contained eighty items "to determine student ability to detect fallacies in reasoning." The readability of the instrument, according to this researcher, ranged from 3.5 to 14.0, and was administered to college freshmen and sophomores. He found no difference for sex, a significant difference between grade levels, a high correlation between general reading ability and "critical reading," and a significant relationship between intelligence and "selected critical reading skills." The particular skills were not given.
Follman (1970) investigated the correlation between critical thinking and critical reading. The Watson-Glaser Critical Thinking Appraisal was used to measure critical thinking ability. In testing the critical reading ability of twelfth grade subjects, Follman used the Reading Comprehension Test (Critical Reading), devised by Martin (1939). Of the fifty items included on this test - five relate to main idea, fifteen to specific details, fifteen are vocabulary items, five test cause and effect, and ten measure inference.

He found a strong correlation between total performance scores, but that individual items on the Watson-Glaser did not consistently discriminate between critical and non-critical readers (Shehell, 1977). However, as Robbins pointed out, the Follman study includes too few subskills considered directly related to critical reading ability.

Sullivan (1977) used a test devised by Maw (1959) to compare "testing strategies in critical reading of good and poor comprehenders." She found no significant relationships between errors made on the STEP Reading Test, Form 3A and the Maw instrument for the two groups. The Maw instrument was intended to measure critical thinking ability and contains single sentence items which are topically unrelated.

Hovland and Weiss (1951) and Whitehead (1971) examined the effect of source credibility. In the Hovland and Weiss
study, four articles were given to two groups. The source for the articles was given high credibility (by the researchers) for one group and low credibility for the other. The high credibility group rated the ideas expressed in the articles as "fair" and changed their opinions on the topic three and one half times more than the low credibility group.

In the Whitehead study, two speeches were given to two groups. Quotes in one speech were attributed to an authority, quotes in the other were simply given as the opinions of the speaker. Whitehead concluded that critical readers were "not impressed" by either speech, whereas "poor" critical readers were "impressed" by the speech which cited the authority. No information was given as to how subjects were measured as being "good" or "poor" critical readers.

In measuring regional differences between subjects, Gadway (1973) considered the abilities to "analyze and reason" and the subjects' "reactions to or opinion about the passages."

All test item examples given in the final report were related to fictional material.

Jacobs and Searfoss (1977) have distinguished and attempted to measure critical reading in their Diagnostic Reading Inventory. There are some questions measuring the subject's ability to draw inferences which could just as easily fit within the designers' inference category of comprehension. However, most of the questions ask for value judgments which are indigenous
to the critical reading of fictional material, which this test contains. Some examples are: "Is stealing from your parents stealing? Explain."; "Who do you feel is right, Woodie, or his parents? Explain."; and "Should Ken have thrown the trap?"

Summary

Critical reading has been variously defined, and its definition appears dependent upon the researcher's purpose in discussing this area of reading. The common denominator of these definitions and the subskills involved is that critical reading is the act of evaluating and passing judgment upon what is read, based upon some known standards or norms.

Researchers in the area of critical reading and reading specialists in general are agreed that this skill is important if one is to function successfully within society. Recognizing the importance of this skill has resulted in attempts to measure its attainment.

However, two conditions exist which lead to confusing and/or conflicting findings. The first of these is directly related to the lack of agreement upon a definition of critical reading and how it relates to the reading process. Is it valid to use a test of critical thinking to measure critical reading? One might assume that critical reading and critical thinking are related. But how strong is the correlation, and
what other factors might be related to critical reading that a
critical thinking appraisal does not measure? What literal
or interpretive skills are necessary for critical reading? Are
they all necessary for critical reading? Are they all necessary,
but not sufficient factors? Are some of them both necessary
and sufficient factors?

The second condition is related to the descriptions of the
studies and the instruments used. Are enough items included
for the measurement of a subskill? Do the items really measure
the subskill they are intended to measure, or are they measuring
some other subskill? Does the researcher recognize that there
are subskills involved? Is the instrument which explains and
defines the terminal behavior measuring the subject's attain­
ment of that skill, or merely his ability to follow directions
and produce that behavior on those particular test items? Is
the critical reading of brief and unrelated test items equiva­
lent to the critical reading of a twenty page chapter in a
textbook? Does the instrument include fictional and/or exposi­
tory test items? Is the critical reader of expository writing
a critical reader of fictional writing? What is the level of
significance? How "strong" is a strong correlation? What was
the strength of other correlational statistics?

To answer these questions and devise adequate measures of
critical reading proficiency, researchers must have complete
descriptions of study instruments, as well as study results, of
prior investigations.
Chapter III

Design of the Study

Purpose

The primary purpose of the present study is to determine the content validity of a researcher-constructed, critical reading assessment. This assessment is an attempt to measure six subskills involved in the critical reading of two topic-related, longer, expository selections. The secondary purpose of the study is to describe the performance of subjects to whom the assessment was administered.

Subjects

The assessment was administered to 224 college students (133 first-year students and 91 third-year students) enrolled in a three-year undergraduate degree program. It was administered during the last two weeks of the school year.

Instrument

This untimed assessment (Appendix B) is comprised of the following: two topic-related, expository selections which are 1,084 and 880 words in length, with the corrected grade level score for each selection being 7th - 8th grade, according to the Dale-Chall formula for estimating readability. The topic of both selections is the behavior of Americans as a group.
There are twenty-nine test items in the following subskill categories (number of items testing each skill follows each category): literal comprehension - 4; recognition of overgeneralization - 4; recognition of factual and opinion/interpretation statements - 9; comparisons - 6; contrasts - 5; and conclusion - 1.

Description of Test Items

The literal comprehension test items are open-ended. Subjects are asked to answer two questions for each selection, the answers to which are stated in the text. Ranking of literal comprehension items from lesser to greater difficulty is: item 1; item 2; item 7; and item 6. The simplicity with which the desired answer was stated in the text was the criterion for assessing difficulty.

The format for the overgeneralization items is as follows. The subject is asked whether or not the first author overgeneralizes. He is then asked, "If you answered "Yes," give two examples from his selection to support your answer." The same procedure is followed in re the second author.

The factual and opinion/interpretation statement items are in the form of quotations taken from the selections. All of the factual statements are taken from the second selection, as there are no factual statements in the first. Three of the opinion/interpretation statements come from the first selection and two
from the second selection. Subjects are asked to judge whether each quotation is: (a) a factual statement that can be proven or disproven, or (b) an opinion/interpretation. The ranking of the opinion/interpretation items from lesser to greater difficulty is: item 10; item 3; item 5; item 13; and item 4. The criterion for ranking these items was the amount of descriptive language used in the statements.

The compare and contrast items are in the following form: subjects are given a list of nine topics "that might be discussed in relation to Americans and their behavior," and asked to circle those topics which "both authors discuss in their analyses of American behavior." Subjects are then asked to discuss the "similarities and/or differences in the authors' views" on each of the topics they have selected. There are five common topics, one of which (materialism) is treated both similarly and differently by the authors. The remaining four topics are viewed differently by the authors. The ranking of contrast items from lesser to greater difficulty is: item 38; item 39; item 37; item 41; and item 40. The criterion for ranking these items was the amount of time the authors spent discussing the topics.

The conclusion item is open-ended. The subject is asked, "Based upon these two selections, what conclusion can the reader come to concerning American character?" and, "What are your reasons for reaching this conclusion?" Four types of conclusions were
anticipated:

Type I - critical - the subject does not accept any of the arguments advanced by either author and decides that no conclusion can be reached by the reader due to lack of substantiation by the authors;

Type II - semi-critical - the subject disregards the lack of substantiation and concludes that: (a) Americans are materialistic, because it is the only characteristic upon which the authors agree; or (b) no conclusion can be reached because the authors disagree on each commonly discussed characteristic;

Type III - non-critical A - the subject concludes that Americans possess all or most of the characteristics discussed by the authors because the authors said Americans possess those characteristics;

Type IV - non-critical B - the subject: (a) disagrees with the authors' characterizations of Americans and characterizes Americans according to his own opinions on the topic, or (b) reaches a conclusion purportedly based upon the selections but which reflects lack of comprehension of the material read.

One point is assigned to each correctly answered item for a total of twenty-nine possible points. A Type I conclusion was considered the only correct response to the conclusion item.
Statistical Analysis

The SPSS (Statistical Package for the Social Sciences) was used to analyze the data. The Oneway Program was used to compare: (1) the mean subskill category performance scores of subjects grouped according to (a) type of conclusion reached, (b) academic performance, and (c) major; and (2) the mean total performance scores of students grouped according to (a) type of conclusion reached, (b) academic performance, and (c) major.

Several _t_-tests have been used to compare (1) the mean difference in a subskill category performance between: (a) first and third year subjects; (b) males and females; and (c) regents and non-regents subjects; (2) the mean total performance scores of subjects grouped according to (a) first and third year subjects, (b) males and females, and (c) regents and non-regents subjects.

The Gutman Scale was used to analyze the difficulty rankings for the literal, opinion/interpretation, and contrast items.

Summary

A researcher-constructed assessment of six subskills involved in the critical reading of two topic-related, longer, expository selections was administered to 224 first and third year college students. The SPSS was used to: (a) determine the content validity of the assessment, and (b) describe the performance of the subjects to whom the assessment was administered.
Chapter IV

Statistical Analysis

Purpose

The primary purpose of the present study is to determine the content validity of a researcher-constructed, critical reading assessment. This assessment is an attempt to measure six subskills involved in the critical reading of two topic-related, longer, expository selections. The secondary purpose of the study is to describe the performance of subjects to whom the assessment was administered.

Findings

The mean score on this twenty-nine point assessment was 18.46 with the range between eight and twenty-eight points. The standard error was .294 and the standard deviation was 4.40.

The percentages of subjects responding correctly or incorrectly to all of the items of each subskill category is illustrated by Figure 1.

Approximately one-half of the subjects correctly identified all the factual statements, with only 1.3 percent incorrectly identifying all of the statements as interpretations or opinions. The next highest category was correct response to literal questions, with 33.5 percent of the subjects responding correctly to all of them. In scoring these items, only complete answers were assigned a point. For
example, the correct response to one of the items was, "colonial and 19th century New England and the South before the Civil War." If a subject omitted "colonial" or "19th century", the answer was considered incorrect.

Figure 1. Percentage of Total Sample Responding Correctly or Incorrectly to All Items of Subskil Categories
The highest category was recognition of overgeneralization, with 31.7 percent of the total sample responding correctly to all four items. However, a nearly equal number of subjects responded incorrectly to all four items. These figures are also important in reference to those subjects who reached a critical conclusion and will be discussed in greater detail later in the chapter.

Twenty-eight percent of the subjects correctly identified all of the opinion/interpretation statements, with 49.1 percent identifying all but one correctly.

The compare and contrast item responses are interesting to consider in relationship to one another. Twenty-four percent of the sample correctly identified all of the topics common to both articles, while only four percent correctly identified all the differences in the authors' discussions of these common topics. Two hundred and seven of the 224 subjects selected materialism as a common topic. Of this number, 94 percent saw the similarity in the discussion and only 20 percent saw how the authors differed in their views on this topic. There is a similar pattern of response for the other correctly identified comparisons as well.

Of the 224 subjects to whom the assessment was administered: 9.4 reached a critical conclusion; 17.0 reached a semi-critical conclusion; 53.9 reached the Type VI non-critical conclusion; and 14.7 reached the Type IV non-critical conclusion. (See
chapter three for a description of these four types of conclusions.)

The correlation between sample performance on each of the subskill categories and (a) total performance score excluding type of conclusion and (b) total performance including the point assigned to a Type I (critical) conclusion if that type of conclusion was reached by the subject is illustrated by Table 1.

Table 1. Pearson's Correlation Between Subskill Category Performance and Total Performance Score:
(a) excluding type of conclusion and (b) including Type I conclusion

<table>
<thead>
<tr>
<th>Subskill Category</th>
<th>(a) excluding type of conclusion</th>
<th>(b) including Type I conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literal questions (4)</td>
<td>.4483</td>
<td>.4489</td>
</tr>
<tr>
<td>Factual statements (4)</td>
<td>.4687</td>
<td>.4698</td>
</tr>
<tr>
<td>Opinion/Interpretation Statements (5)</td>
<td>.5014</td>
<td>.4994</td>
</tr>
<tr>
<td>Overgeneralizations (4)</td>
<td>.6955</td>
<td>.7053</td>
</tr>
<tr>
<td>Comparisons (6)</td>
<td>.5320</td>
<td>.5267</td>
</tr>
<tr>
<td>Contrasts (5)</td>
<td>.7313</td>
<td>.7830</td>
</tr>
</tbody>
</table>

n < .001 in each cell

A significant correlation is to be expected between the whole and its parts. What is interesting to note is the
decrease in correlation in the second column for all subskill categories except overgeneralization and contrast. For these two categories there is an increase in correlation. The correlations between overgeneralization and column 1 and column 2 are also high considering the fact that this category contains the same number of items as the literal and factual categories and fewer items than the categories of opinion/interpretation, comparison, and contrast.

Comparison of Subjects by Type of Conclusion Reached

Analysis of variance was used to compare the mean subskill category performance scores of subjects grouped according to type of conclusion reached. Table 2 shows the results for the literal questions.

The variance in mean scores was significant at the .005 level. The maximum number of correct responses was four for all groups. The minimum number of correct responses was two for the critical conclusion group, one for the semi-critical conclusion group, and zero for the two non-critical groups. However, the semi-critical group has a slightly higher mean score than the critical group. There is an overlap in the confidence intervals for critical and semi-critical groups, but no overlap between these first two groups and the non-critical groups.
Table 2. Analysis of Variance of Means on Literal Questions by Conclusion Type Groups

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>3</td>
<td>14.23</td>
<td>4.74</td>
<td>4.48</td>
<td>&lt;.005</td>
</tr>
<tr>
<td>Within groups</td>
<td>220</td>
<td>232.80</td>
<td>1.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>223</td>
<td>247.03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Range</th>
<th>95% Confidence Interval for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I (critical)</td>
<td>3.29</td>
<td>2 - 4</td>
<td>2.90 to 3.67</td>
</tr>
<tr>
<td>Type II (semi-critical)</td>
<td>3.34</td>
<td>1 - 4</td>
<td>3.06 to 3.62</td>
</tr>
<tr>
<td>Type III (non-critical A)</td>
<td>2.78</td>
<td>0 - 4</td>
<td>2.59 to 2.97</td>
</tr>
<tr>
<td>Type IV (non-critical B)</td>
<td>2.67</td>
<td>0 - 4</td>
<td>2.27 to 3.06</td>
</tr>
</tbody>
</table>

The mean differences between the four groups on the number of factual statements correctly identified is illustrated by Table 3.

The level of significance for this subskill category was <.001. For identification of factual statements, the semi-critical group again has a higher mean score than the critical group. The minimum number of correct responses is also higher for the semi-critical group, with an accompanying overlap of confidence intervals for these two groups. However, there is
no overlap between these first two groups and the non-critical groups.

Table 3. Analysis of Variance of Means for Identification of Factual Statements by Conclusion Type Groups

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>3</td>
<td>8.65</td>
<td>2.85</td>
<td>5.38</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Within groups</td>
<td>220</td>
<td>117.91</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>223</td>
<td>126.55</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Range</th>
<th>95% Confidence Interval for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I (critical)</td>
<td>3.6</td>
<td>2 - 4</td>
<td>3.26 to 3.33</td>
</tr>
<tr>
<td>Type II (semi-critical)</td>
<td>3.8</td>
<td>3 - 4</td>
<td>3.69 to 3.90</td>
</tr>
<tr>
<td>Type III (non-critical A)</td>
<td>3.3</td>
<td>1 - 4</td>
<td>3.20 to 3.46</td>
</tr>
<tr>
<td>Type IV (non-critical B)</td>
<td>3.2</td>
<td>1 - 4</td>
<td>2.91 to 3.53</td>
</tr>
</tbody>
</table>

A comparison of the four types of conclusion groups on number of opinion/interpretation statements correctly identified is shown in Table 4:

For this subskill the critical group has both the higher mean score and the higher minimum number of correct responses. Once again the confidence intervals overlap between the first two groups, but not between these groups and the non-critical
groups. The level of significance for this subskill category performance was < .002.

Table 4. Analysis of Variance of Means for Identification of Opinion/Interpretation Statements by Conclusion Type Groups

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>3</td>
<td>12.86</td>
<td>4.29</td>
<td>5.23</td>
<td>&lt; .002</td>
</tr>
<tr>
<td>Within groups</td>
<td>220</td>
<td>180.50</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>223</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Range</th>
<th>95% Confidence Interval for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I (critical)</td>
<td>4.4</td>
<td>3 - 5</td>
<td>4.16 to 4.70</td>
</tr>
<tr>
<td>Type II (semi-critical)</td>
<td>4.3</td>
<td>2 - 5</td>
<td>4.06 to 4.52</td>
</tr>
<tr>
<td>Type III (non-critical A)</td>
<td>3.3</td>
<td>1 - 5</td>
<td>3.66 to 3.99</td>
</tr>
<tr>
<td>Type IV (non-critical B)</td>
<td>3.7</td>
<td>1 - 5</td>
<td>3.35 to 4.41</td>
</tr>
</tbody>
</table>

The difference in mean performance for the subskill of recognition of overgeneralization are illustrated in Table 5.

Here the differences between groups are more dramatic and are significant at the < .001 level. Not only does the critical group have the highest mean score and highest minimum number of correct responses, there is also no overlap in confidence intervals between this group and the semi-critical group.
Table 5. Analysis of Variance of Means for Recognition of Overgeneralizations by Conclusion Type Groups

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>3</td>
<td>145.58</td>
<td>48.53</td>
<td>25.55</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Within groups</td>
<td>220</td>
<td>417.77</td>
<td>1.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>223</td>
<td>563.36</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Range</th>
<th>95% Confidence Interval for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I (critical)</td>
<td>3.90</td>
<td>2 - 4</td>
<td>3.71 to 4.10</td>
</tr>
<tr>
<td>Type II (semi-critical)</td>
<td>3.30</td>
<td>0 - 4</td>
<td>2.53 to 3.47</td>
</tr>
<tr>
<td>Type III (non-critical A)</td>
<td>1.66</td>
<td>0 - 4</td>
<td>1.40 to 1.92</td>
</tr>
<tr>
<td>Type IV (non-critical B)</td>
<td>1.33</td>
<td>0 - 4</td>
<td>.86 to 1.31</td>
</tr>
</tbody>
</table>

Table 6 shows the mean differences between groups on ability to note comparisons between the two articles.

The probability frequency, although significant ($p < .02$), reflects the relatively even performance of all groups on this task. Again there is an overlap in confidence for the critical and semi-critical groups, with the semi-critical groups having the higher mean score and higher minimum number of correct responses. As has been the case previously, the non-critical groups, although they overlap with one another, do not overlap with the first two groups.
Table 6. Analysis of Variance of Means for Recognition of Comparisons by Conclusion Type Groups

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>3</td>
<td>15.81</td>
<td>5.27</td>
<td>3.31</td>
<td>&lt;.02</td>
</tr>
<tr>
<td>Within groups</td>
<td>220</td>
<td>350.17</td>
<td>1.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>223</td>
<td>365.98</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Range</th>
<th>95% Confidence Interval for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>4.86</td>
<td>2 - 6</td>
<td>4.37 to 5.34</td>
</tr>
<tr>
<td>(critical)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type II</td>
<td>4.97</td>
<td>3 - 6</td>
<td>4.63 to 5.27</td>
</tr>
<tr>
<td>(semi-critical)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type III</td>
<td>4.33</td>
<td>1 - 6</td>
<td>4.10 to 4.56</td>
</tr>
<tr>
<td>(non-critical A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type IV</td>
<td>4.36</td>
<td>0 - 6</td>
<td>3.90 to 4.93</td>
</tr>
<tr>
<td>(non-critical B)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Performance of subjects on the contrast recognition items is illustrated in Table 7.

The minimum number of correct responses was zero for all four groups. And, although the semi-critical group has the highest mean, the means for this group and the critical group are considerably higher than the means for the two non-critical groups with the level of significance being <.001.
Table 7. Analysis of Variance of Means for Recognition of Contrasts by Conclusion Type Groups

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>3</td>
<td>123.72</td>
<td>41.24</td>
<td>23.11</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Within groups</td>
<td>220</td>
<td>322.72</td>
<td>1.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>223</td>
<td>446.50</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Group       | Mean | Range  | 95% Confidence Interval for Mean
----------|------|--------|---------------------------------|
Type I      | 2.76 | 0 - 4  | 2.17 to 3.35                    |
  (critical) |      |        |                                 |
Type II     | 2.92 | 0 - 5  | 2.52 to 3.32                    |
  (semi-critical) |      |        |                                 |
Type III    | 1.16 | 0 - 4  | .96 to 1.36                     |
  (non-critical A) |      |        |                                 |
Type IV     | 1.27 | 0 - 4  | .73 to 1.76                     |
  (non-critical B) |      |        |                                 |

Analysis of variance was also used to compare the mean total performance scores for subjects grouped according to type of conclusion reached. These means, of necessity, exclude the point assigned for a critical conclusion. Table 8 illustrates the total performance for these four groups.

The difference between the non-critical groups and the first two groups is most obvious in this comparison. The mean total performance for the critical group was 22.67. The semi-critical group has a mean score of 22.37. The mean total performance scores for the non-critical groups was 16.97.
and 16.61. As would be expected from the subskill means, there is an overlap of confidence intervals for the critical and semi-critical groups and no overlap between these two groups and the non-critical groups. The minimum number of correct responses for the critical group was seventeen, and only fourteen for the semi-critical group. However, it should be noted that the only subject to obtain a perfect score on this measure was in the semi-critical group.

Table 8. Analysis of Variance of Mean Total Performances by Conclusion Type Groups

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>3</td>
<td>1356.72</td>
<td>452.24</td>
<td>36.27</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Within groups</td>
<td>220</td>
<td>2743.26</td>
<td>12.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>223</td>
<td>4099.98</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Range</th>
<th>95% Confidence Interval for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I (critical)</td>
<td>22.67</td>
<td>17 - 27</td>
<td>21.44 to 23.90</td>
</tr>
<tr>
<td>Type II (semi-critical)</td>
<td>22.37</td>
<td>14 - 23</td>
<td>21.33 to 23.41</td>
</tr>
<tr>
<td>Type III (non-critical A)</td>
<td>16.97</td>
<td>9 - 25</td>
<td>16.35 to 17.59</td>
</tr>
<tr>
<td>Type IV (non-critical B)</td>
<td>16.61</td>
<td>8 - 26</td>
<td>15.16 to 18.05</td>
</tr>
</tbody>
</table>

Based upon examination of the total sample performance and comparison of the subjects by type of conclusion reached, the
ability to recognize overgeneralizations was an important factor in reaching a critical conclusion about two topic-related, expository selections. However, one subject in the critical conclusion group recognized specific overgeneralizations in one selection, but not in the other. Also, some subjects in the semi-critical and non-critical conclusion groups recognized the overgeneralizations in both selections.

Thus, the ability to recognize overgeneralizations does not by itself determine type of conclusion reached on this assessment. It does, however, have a strong influence. No subject who failed to recognize any of the overgeneralizations reached a critical conclusion; and 95 percent of those subjects who reached a critical conclusion correctly identified all four overgeneralizations asked for.

The other important factors appear to be recognition of opinion/interpretation statements and the ability to see contrasting views. But again, these are not sufficient to explain the difference in type of conclusion reached. Some subjects who scored as well as, or better than, the critical conclusion subjects on these measures did not reach the critical conclusion.

**Difficulty Rankings of Literal, Opinion/Interpretation and Contrast Items**

The Gutman Scale was used to analyze the difficulty rankings for the literal, opinion/interpretation, and contrast items.
Ranking of literal comprehension items from greatest to least difficulty was item 6, item 7, item 2, and item 1. The linguistic simplicity with which the answer was stated in the texts was the criterion for assessing the difficulty of these items. One hundred and sixteen of the 224 subjects gave the correct response to item 6, 165 gave the correct response to item 7, 197 gave the correct response to item 2, and 170 gave the correct response to item 1.

Thus, the difficulty ranking was as anticipated with the exception of item 1. Item 1 was the literal question referred to earlier in this chapter as an example of the requirement that a subject give a complete answer in order to receive the point assigned to an item. Of the 54 subjects who did not receive credit for this item, 33 gave incomplete responses.

The coefficient of reproducibility was .84. That is, there is an 84 percent probability that the response pattern would be the same if this group of questions were posed to future subjects.

The ranking of the opinion/interpretation items from greatest to least difficulty was item 4, item 13, item 5, item 3, and item 10. The criterion for ranking these items was the amount of descriptive language used in the statements. Eighty-three respondents answered item 4 correctly, 190 answered item 13 correctly, 203 answered item 5 correctly, 203 answered item 3 correctly, and 209 answered item 10 correctly. The coefficient of reproducibility for these items was .92.
The ranking for the contrast items from greatest to least difficulty was item 40, item 41, item 37, item 39, and item 33. Seventy-eight subjects responded correctly to item 40, 35 subjects responded correctly to item 41, 53 subjects responded correctly to item 37, 79 subjects responded correctly to item 39, and 121 subjects responded correctly to item 38. The criterion for ranking these items was the amount of time the authors spent discussing the topics. A possible explanation for the low correct response rate to item 41 could be the fact that Aldridge mentions both positive and negative attitudes toward the home and Commager refers only to the American's carelessness toward the home. The coefficient of reproducibility for these items was .79.

Academic Performance and Assessment Performance

For this analysis of variance, the score used to indicate a subject's academic performance was his college grade point average (GPA). The grouping of subjects according to GPA and the number of subjects in each group are as follows:

- Group 1 - 1.4 or below (4)
- Group 2 - 1.5 - 1.9 (10)
- Group 3 - 2.0 - 2.4 (31)
- Group 4 - 2.5 - 2.9 (60)
- Group 5 - 3.0 - 3.4 (50)
- Group 6 - 3.5 or above (33)
Grade point averages were not available for 31 subjects.

The three top performing groups for each subskill category and total performance excluding type of conclusion are shown in Table 9. A cursory inspection of this table reveals a very close relationship between academic performance and assessment performance.

The three groups with the highest mean total performance scores were Group 6 with a mean score of 20.6, Group 5 with a mean score of 19.4, and Group 4 with a mean score of 17.6. Group 1 had the lowest mean score — 15.5.

The highest scoring three groups for the categories of factual, overgeneralization, and contrast were again, in descending order, Group 6, Group 5, and Group 4.

The level of significance for the comparison items was < .01. In this category, Group 2 was highest, followed by Groups 5 and 6.

For the category of opinion/interpretation, the difference in means scores was not statistically significant. Group 1 was the highest scoring group, followed by Group 5 and 6.

Differences between groups were not statistically significant for the literal items. It should be noted, however, that Group 1 subjects had the highest mean for this category.
<table>
<thead>
<tr>
<th>Group</th>
<th>Total Performance</th>
<th>Factual</th>
<th>Overgeneralization</th>
<th>Contrast</th>
<th>Comparison</th>
<th>Opin/Int.</th>
<th>Literal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>17.6</td>
<td>3.50</td>
<td>1.77</td>
<td>1.32</td>
<td></td>
<td>4.50</td>
<td>3.75</td>
</tr>
<tr>
<td>(1.4 or below)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>19.4</td>
<td>3.62</td>
<td>2.20</td>
<td>1.82</td>
<td>4.76</td>
<td>4.1</td>
<td>3.00</td>
</tr>
<tr>
<td>(1.5 - 1.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>20.6</td>
<td>3.55</td>
<td>2.39</td>
<td>2.32</td>
<td>4.71</td>
<td>4.11</td>
<td>3.00</td>
</tr>
<tr>
<td>(2.0 - 2.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.5 - 2.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3.0 - 3.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3.5 or above)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Level of Significance: $p < .001$, $p < .05$, $p < .001$, $p < .001$, $p < .01$, $p = n.s.$, $p = n.s.$

*Numbers in parentheses represent Grade Point Average (GPA) for each group.*
The results of a cross-tabulation of types of conclusions reached by academic performance groups is shown in Table 10.

The outstanding observation to be made from this Table of the subjects in the lowest two academic performance groups reached either a critical or semi-critical conclusion. Group 6 had the highest percentage of subjects reaching a critical conclusion and Group 5 had the second highest percentage of subjects reaching a critical conclusion. It would appear that, once again, the ability to recognize overgeneralizations and contrasts is related to reaching a critical conclusion on this assessment. These differences in types of conclusions reached were significant at the <.05 level.
Table 10. Academic Performance & Type of Conclusion Reached

<table>
<thead>
<tr>
<th>Conclusion</th>
<th>Group 1 (1.4 or below)</th>
<th>Group 2 (1.5-1.9)</th>
<th>Group 3 (2.0-2.4)</th>
<th>Group 4 (2.5-2.9)</th>
<th>Group 5 (3.0-3.4)</th>
<th>Group 6 (3.5 or above)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I critical</td>
<td>0%</td>
<td>0%</td>
<td>3%</td>
<td>5%</td>
<td>14%</td>
<td>16%</td>
</tr>
<tr>
<td>Type II semi-critical</td>
<td>0%</td>
<td>0%</td>
<td>7%</td>
<td>23%</td>
<td>20%</td>
<td>21%</td>
</tr>
<tr>
<td>Type III non-critical A</td>
<td>75%</td>
<td>50%</td>
<td>74%</td>
<td>63%</td>
<td>50%</td>
<td>53%</td>
</tr>
<tr>
<td>Type IV non-critical B</td>
<td>25%</td>
<td>50%</td>
<td>16%</td>
<td>8%</td>
<td>16%</td>
<td>10%</td>
</tr>
</tbody>
</table>

n < .03
Type of New York State Diploma and Assessment Performance

The difference in academic performance between regents and non-regents high school graduates in this study was not statistically significant. However, 50 percent of the regents subjects had a GPA of 3.0 or higher as compared with 34 percent of the non-regents students.

In comparing mean subskill category performance scores, the regents group did significantly better than the non-regents group in all categories except comparison. However, the regents mean score in this category was higher than the non-regents mean score. The difference in type of conclusion reached was statistically significant at the <.01 level. Eleven percent of the regents group reached a critical conclusion, as compared with the non-regents percentage of five. Twenty-two percent of the regents group came to a semi-critical conclusion, while only seven percent of the non-regents group reached this type of conclusion.

Again, the greatest statistical differences were found when comparing performances on overgeneralization (p <.01) and contrast (p <.001). Total performance mean score, excluding type of conclusion was significant at the <.001 level. The mean total for the regents group was 19.4. The mean total for the non-regents group was 16.4.

College Major and Assessment Performance

For this analysis of variance between mean scores, sub-
Subjects were grouped according to college divisions. For example, fine arts included subjects majoring in art, dance, music, and theatre. Because of the proportionately high number of business majors, these subjects were considered a separate group rather than part of the social sciences division. These groups and the number of subjects in each group is as follows:

- undeclared (22)
- business (37)
- natural and mathematical sciences (49)
- social sciences (33)
- humanities (27)
- fine arts (24)
- human services (16)
- physical education and recreation (16)

There was no statistically significant differences among groups on academic performance (GPA).

The three top performing groups for each subskill category and total performance, excluding type of conclusion, are shown in Table 11.

When these eight groups were compared on total performance, excluding type of conclusion, the level of significance was .003. The top three groups were humanities with a mean score of 21.1, natural and mathematical sciences with a mean score of 19.2, and human services with a mean score of 13.6.
<table>
<thead>
<tr>
<th></th>
<th>Total Performance</th>
<th>Contrast</th>
<th>Opin/Int.</th>
<th>Factual</th>
<th>Comparison</th>
<th>Literal</th>
<th>Overgeneralization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undeclared</td>
<td></td>
<td></td>
<td>3.6</td>
<td></td>
<td>3.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.4</td>
</tr>
<tr>
<td>Natural &amp; Mathematical Sciences</td>
<td>19.2</td>
<td>1.3</td>
<td>4.1</td>
<td>3.6</td>
<td>4.6</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>Social Sciences</td>
<td></td>
<td></td>
<td>4.0</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humanities</td>
<td>21.2</td>
<td>2.3</td>
<td>4.5</td>
<td>4.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine Arts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.9</td>
</tr>
<tr>
<td>Human Services</td>
<td>18.6</td>
<td>1.9</td>
<td></td>
<td></td>
<td>4.8</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td>Physical Education &amp; Recreation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of Significance</td>
<td>$p &lt; .003$</td>
<td>$p &lt; .02$</td>
<td>$p &lt; .003$</td>
<td>$p &lt; .03$</td>
<td>$p=n.s.$</td>
<td>$p=n.s.$</td>
<td>$p=n.s.$</td>
</tr>
</tbody>
</table>
There was a significant difference among groups (p<.02) on the contrast category performance. The top three groups were humanities with a mean score of 2.3, human services with a mean score of 1.9, and natural and mathematical sciences with a mean score of 1.3.

There was a significant difference among groups on opinion/interpretation category performance (p<.003). The top three groups were humanities, natural and mathematical sciences, and social sciences. This was the only category for which there was no overlap in the 95 percent confidence intervals for the means.

The top three groups for the factual category (p<.03) were natural and mathematical sciences and subjects with an undeclared major and equal means of 3.6 and social sciences with a mean of 3.5.

There were no statistically significant differences in performance among groups for the categories of comparison, literal, and overgeneralization. However, humanities had the highest mean for each category.

When compared by type of conclusion reached, cross-tabulation indicated there were no statistically significant differences among groups. However, this comparison of percentages of subjects reaching each type of conclusion indicates humanities majors to have the greatest number of subjects reaching critical or semi-critical conclusions. Forty-four percent of these subjects reached a critical or semi-critical conclusion. The percentage for natural and mathematical sciences majors was
thirty-four, and the percentage for human services majors was thirty-one.

The importance of these percentages is apparent when one compares these three groups in re their performances on the overgeneralization and contrast items. Humanities scored highest in both of these subskill categories. Human services was among the top three groups for contrasts but not overgeneralizations. Thus, it appears that the subskill categories of overgeneralization and contrast are again related to type of conclusion reached.

Sex and Assessment Performance

There were 93 males and 126 females in the sample and a t-test was used to compare means between the two groups. There was no statistically significant difference between the two groups in academic performance.

There were no significant differences between the sexes in performance in the categories of factual, opinion/interpretation, comparison, and overgeneralization. There was also no significant difference for total performance score excluding type of conclusion. Males scored slightly higher than females on factual, opinion/interpretation, and overgeneralization items. Females scored slightly higher on comparison and total performance excluding type of conclusion. Females scored significantly higher (p < .02) on the contrast items, and also on the literal items (p < .005).
Level of significance on a cross-tabulation for type of conclusion reached is not considered statistically significant but worth noting ($p=0.09$). Ten percent of the females and eight percent of the males reached a critical conclusion. Nineteen percent of the females and fourteen percent of the males reached a semi-critical conclusion.

**Grade Level and Assessment Performance**

A $t$-test was used to compare mean performances of the 133 first year college students with the 91 third year college students in the sample. There was no significant difference in the academic performance of these two groups.

Third year students scored higher than first year students on all subskill categories. They scored significantly higher on comparison ($p<0.04$) and overgeneralization ($p<0.001$). The conclusion item is most interesting. Ten percent of the first year subjects reached a critical conclusion. Seven percent of the third year students reached a critical conclusion. This would appear contradictory to previous findings concerning the relationship between type of conclusion reached and overgeneralization and contrast category performance. However, 22 percent of the third year students came to a semi-critical conclusion, while 13 percent of the first year students reached this type of conclusion.
Summary

There was a normal distribution of scores on the assessment. Pearson's correlation showed the appropriate significant relationship between the subskill categories and the total performance score. Guttman scaling supported the difficulty ranking for the opinion/interpretation items. It did not support the rankings for the literal and contrast items.

Sample subjects were grouped according to the variables of academic performance, type of high school diploma, major, sex, and grade level. Several relationships were found among these variables and assessment performance. Those subjects with the highest GPA scored significantly higher than other subjects. Regents diploma subjects scored significantly higher than non-regents subjects. Humanities majors scored significantly higher than other majors. There was no significant performance between the sexes. Third year subjects scored significantly higher than first year subjects.

With the exception of the variable of sex, total performance by the different groups appeared consistently related to performance on the subskills of overgeneralization and contrast. Pearson's correlation for the entire sample confirmed this. Furthermore, there was a strong relationship between performance on these two categories and type of conclusion reached. Analysis of these groups and the total sample grouped according to type of conclusion reached showed that those subjects who performed

54
well on the overgeneralization and contrast items reached a
critical or semi-critical conclusion. Those subjects who
scored low on both of these categories did not reach these types
of conclusions. However, high performance in these two cate-
gories did not always lead to a critical conclusion rather
than a semi-critical conclusion.
Chapter V

Conclusions and Implications

Purpose

The primary purpose of the present study is to determine the content validity of a researcher-constructed, critical reading assessment. This assessment is an attempt to measure six subskills involved in the critical reading of two, topic-related, longer, expository selections. The secondary purpose of the study is to describe the performance of subjects to whom the assessment was administered.

Conclusions

Five questions were asked at the onset of this study. The first was Does the assessment produce a normal range of variability in the total performance scores? The range was normal with a skewness of +.012.

The second question was Is there a relationship between the type of conclusion reached by subjects and their performance scores for other subskill categories? It was the assumption of this study that in order to critically read any written material, the reader must first literally comprehend what has been read. Furthermore, with the two topic-related, expository selections included in this assessment, the reader must compare and contrast the authors' views as well as recognize facts, opinions/interpretations and overgeneralizations, and reach a
conclusion about the topic of discussion based upon all arguments presented. That is, the reader would not be able to reach a critical conclusion without recognition of the strengths and weaknesses of the authors' discussions.

The two selections in this assessment were included because of their contrasting views of American behavior and their propensities to overgeneralize and state opinion/interpretations without substantiating factual evidence. Indeed, the first selection contains no facts throughout. The second selection does contain verifiable statements, but these often relate to governmental policy rather than the actions of the American populace which the author purports to be discussing. The observations of both authors are generalized to the entire American population with no allowance made for individual differences or exceptions.

Subjects who reached critical or semi-critical conclusions did significantly better in all other subskill categories than the two types of non-critical readers. Of particular note was performance on recognition of overgeneralizations. Twenty of the twenty-one subjects who reached a critical conclusion noted all four of the required overgeneralizations. The twenty-first saw the overgeneralization in one selection but not the other. However, it was found that while high performance on the subskill categories of literal comprehension, recognition of factual statements, opinion/interpretations, and the abilities to compare,
contrast of all overgeneralizations was necessary to reach a critical conclusion, it was not sufficient. The one subject who had correct responses to all items, excluding type of conclusion, reached a semi-critical conclusion.

There are two possible explanations for this. The first is that students may be reluctant to pass personal judgment upon what they read, especially in a test situation. There may be reticence to proclaim such material as worthless even when all the flaws are recognized.

The second possible explanation for failure to reject this unsubstantiated expository material may be that the wording of the conclusion question implies that a positive conclusion should be reached. The question reads Based upon these two selections, what conclusion can the reader come to concerning American Character? All subjects may not realize that by stating that no conclusion can be reached concerning American Character, they have indeed reached a conclusion on this topic based upon the selections. Of the possible explanations for response to this item, the latter appears as plausible as the former, and further consideration should be given to the wording of this item.

The third question posed by the present study was Is there a relationship between performance on more difficult items and performance on items of less difficulty for the skill categories of literal comprehension, recognition of contrasts, and opinion/
interpretation recognition? Although the coefficients of reproducibility were reasonably high, none of these reached the desired coefficient of .95. This would indicate that the scaled items varied in difficulty for reasons other than, or in addition to, the stated criteria (and/or the wording of the questions, in the case of the literal comprehension questions). The criterion for each set of items should be re-examined.

The fourth and fifth questions dealt with total and subskill category performance scores of the sample grouped according to academic performance, type of New York State diploma, major, grade level, and sex. Subjects with the higher GPAs scored significantly higher on the assessment, and none of the subjects with a GPA of 1.4 - 1.9 reached a critical or semi-critical conclusion.

Regents diploma subjects also scored significantly higher than non-regents subjects. Fifty percent of the regents students had a GPA of 3.0 or higher, as opposed to 31% of the non-regents subjects. Hence, there appears to be a strong relationship between academic performance and critical reading ability as measured by this assessment. If IQ is a predictor of academic performance, then this finding supports previous research in this area.

Humanities majors scored significantly higher than other majors on total performance. And third year subjects scored significantly higher than first year subjects. The latter
finding is in agreement with that of Gall (1973), Johnson (1974), and Downing (1976).

There was no significant difference between sexes in total performance although females scored slightly higher. Females did significantly better on the contrast and literal comprehension items and had a higher percentage reaching critical and semi-critical conclusions. Previous research also indicated no significant difference between these groups in total performance.

When subjects were grouped by major, grade level, and sex, there were no significant differences between subjects on academic performance. Therefore, in addition to academic performance, number of years in college and major also appear related to critical reading ability as measured by this assessment.

Total performance by the different groupings appeared consistently related to performance on the subskill categories of overgeneralization and contrast. And, as discussed earlier, there was a consistent relationship between performance on these two categories and type of conclusion reached.

Implications for Research

Future research in the measurement of critical reading ability should, as stated earlier, be confined to the measurement of specific subskills related to the critical reading of a specific type of writing. The present study attempted to measure those subskills thought necessary for the critical reading of two topic-related, expository selections. In this
study, general critical reading ability (as indicated by total performance score) and the ability to reach a critical conclusion appeared strongly related to the recognition of overgeneralizations and contrasts, and to a lesser extent, recognition of opinions/interpretations. To conclude that those subjects who performed well on this assessment are critical readers, and those who did not are not critical readers, would be invalid. They are critical, semi-critical, or non-critical readers of the type of material included in the assessment as measured by the subskill items included on this assessment.

Furthermore, the conclusion that subjects in this study are or are not critical readers of this type of material is only as valid as the assessment used to measure this ability. Replication of this pilot study should not be attempted without consideration for several possible refinements of the assessment:

1. revision of the conclusion item;
2. change in the format of the comparison subskill category -- as it now stands, it is possible for a subject to circle all the choices and receive five of the six possible points in this category; and
3. further investigation into why some items thought to be less difficult than others (based upon previously stated criteria) were apparently more difficult.

Implications for Classroom Practice

This assessment would indicate that more instructional time
should be spent on the development of the subskills of recognition of opinions/interpretations and overgeneralizations in individual expository selections. Of equal importance, the student should be trained to relate what is presently being read to what was previously known. To promote this interrelating of information, the subskills of recognition of similar and contrasting views should be emphasized.

Pertinent to these areas of instruction are two other areas of apparent weakness. First, students appear to be able to identify factual statements, but not judge whether or not those statements sufficiently substantiate an author's opinion/interpretation. Second, many students who appear able to recognize flaws in an argument, are unwilling to evaluate the validity of the total argument based upon their recognition of these flaws. That is, they see the lack of substantiation for a presented point of view but will still consider that point of view worthy of reiteration rather than rejection.

The student should be aware that the purpose of analyzing an article or book is to reach a personal critical conclusion in re that article or book -- to determine whether the selection expands his knowledge of a topic, or should leave that knowledge justifiably unaltered.
References


Cameron, J.R. Read critically or join the mob. Journal of Reading, October 1968, 15, 24-26.


Harris, A. How to increase reading ability (6th ed.). New York: David McKay Company, 1975, 484.


Nancy, E.S. Literal and critical reading in science. Journal of Experimental Education, 1953, 27, 57-64.


Pioronek, F.T. Acquisition of specific critical reading skills and development of learning style in an individualized reading program and a basal reading program (Doctoral dissertation, Boston University School of Education, 1974). Dissertation Abstracts International, 1974, 35, 1A-1292A.

Queen, B. Black and white and read all over. Reading Quarterly, July 1973, 6 (3), 59-76.


Smith, R.J. Evaluating the teaching of higher level comprehension skills. Paper presented at University of Wisconsin seminar, Madison, May 1972. (ERIC Document Reproduction Service No. ED 065 830)


Woodruff, A.D. Human behavior and the acquisition of competence in critical reading. Paper presented at the meeting of the International Reading Association Preconvention Institute, Boston, April 1968. (ERIC Document Reproduction Service No. ED 029 759)
Appendix A

Extended Lists of Critical Reading Skills

Raciti's (1967) Specific Skills

1. Draw conclusions
2. Predict outcomes
3. Draw inferences
4. Recognize cause and effect
5. Make comparisons
6. Distinguish between fact and fancy, relevant and irrelevant information, and similarities and differences
7. Judge skill of author in writing
8. Accept or reject author's facts
9. Understand need to suspend judgment until more information is known
10. Be willing to accept the ideas of the author if he is an expert
11. Be able to judge the bias of the author

Wolf, Huck, and King's (1968) List of Skills

General Abilities
1. Recognize that printed material is just one source of information
2. Comprehend various reading materials that present different viewpoints on a topic
3. Keep reading until one has all the information necessary to come to a conclusion

Specific Abilities
1. Recognize the author's point of view and purpose
2. Recognize the publisher's point of view and purpose
3. Determine the expertise of the author
4. Distinguish between fact and opinion
5. Follow the sequence presentation and judge its logic
6. Compare and contrast and judge appropriateness to increasing the reader's knowledge
7. Form an opinion, relating to past experience, and realizing when one does not have enough experience to form an opinion
8. Locate and select relevant materials
9. Recognize when important facts have been omitted
10. Identify persuasion techniques such as:
    a. emotional appeal
    b. glittering generalities
    c. endorsements
    d. inferring relationships that don't exist
11. Determine the accuracy of graphic presentations
12. Identify literary form (such as fiction, non-fiction and satire)
13. Come to a personal opinion of the quality of literature through analysis of such elements as:
   a. theme
   b. character development
   c. style

Williams' (1959) List of Skills

1. Anticipating outcomes
2. Appreciating humor, plot
3. Classifying ideas
4. Comparing and contrasting
5. Critical thinking
6. Distinguishing fact and fancy
7. Distinguishing fact and opinion
8. Drawing conclusions
9. Establishing sequence
10. Establishing cause and effect
11. Evaluating author's attitude
12. Evaluating and reacting to ideas in light of the author's purpose
13. Evaluating and solving problems
14. Evaluating summaries
15. Finding information to prove or disprove a statement
16. Forming an opinion
17. Forming sensory impressions
18. Generalizing
19. Identifying elements of style
20. Identifying and evaluating character traits
21. Interpreting figurative and idiomatic language
22. Interpreting ideas implied, not stated
23. Judging author's statements on background of author
24. Judging reasonableness and relevance
25. Making inferences
26. Making judgments
27. Perceiving relationships
28. Predicting outcomes
29. Reacting to the mood or tone of a selection
30. Recognizing emotional reactions and motives
31. Recognizing story problems and plot structure
32. Relating story experiences to personal experiences
33. Research
Appendix B

Assessment

(Item numbers for the Gutman Scale analysis correspond with question numbers except for the contrast items. Contrast item numbers have been inserted for the convenience of the reader and did not appear on the administered assessment.)

NAME _______________________________________

AGE _______________________________________

MAJOR _______________________________________

MINOR _______________________________________

Were you a college student anywhere else before coming to the Alternate College?

If so, WHERE? ___________________________ and FOR HOW LONG? ____________

Circle the level accurate for you as an A.C. student:

Level I - 1st semester    Level III - 1st semester
Level I - 2nd semester    Level III - 2nd semester

High School Diploma (circle appropriate one): Regents  Non-Regents

_____________________________  ________________________________

AMERICANS AND CULTURE

   John W. Aldridge

The plain fact is that the basis for the civilized life - separate from the intellectual and artistic life - has never existed in this country except at certain odd times and places in the past. Those times and places were colonial and nineteenth century New England and the South before the Civil War. But the flowering of high culture in these areas was probably too
dependent on temporary regional conditions and declining European influences. Or it was not strong enough to alter the utilitarian nature of the way of life which was spreading westward and would shortly fix the character of our national existence.

By the time most immigrants came to the country, the ideals of political and religious freedom had been largely replaced by a promise of unlimited opportunity to improve the material conditions of life. To be free not only to worship and vote as one wished, but to become as rich and powerful as one's talents for using people and materials would permit. Presumably, there was some slight suggestion in the immigrants' agreement with America that if material conditions could be improved enough, the luxuries and graces would follow in the natural course of things. But the pioneers were not, for the most part, cultured men. And whatever vision of civilization they may have brought with them from Europe could not survive long in the hard life of the frontier.

The physical effort needed to tame the wilderness and provide the things necessary just to live was such that no thought could be given to the problem of making life beautiful.

But then, right at that point, something in our evolutionary mechanism went very wrong, and has remained very wrong ever since. Logically, we should have evolved to the next phase of growth. With the close of the frontier and the spread of material abundance, a demand should have arisen for luxury and the cultivation of taste, intellect and the social pleasantries. Instead, life
in America became frozen - apparently for good - at the utilitarian level. We simply went on providing more and more conveniences for greater and greater numbers of people. The hard work and cleverness which had formerly gone into pushing back the physical frontier was used to push back the "material things" frontier.

Nearly everything we have created here - except in the old areas of the original colonies - is stamped with the brand of the prairie and the frontier settlement. The typical main street of Anywhere - Nowhere, U.S.A. is still basically that of a Montana cowtown in the 1880's. It all has the appearance of having been thrown together in a great hurry to provide temporary shelter and the bare needs of life for a people who will be moving on to a new home. We take it for granted that new houses will not be roomy or well-built or pleasant to look upon. We assume they will offer no luxury or ease beyond that provided by the average motel, and that we can litter the countryside around them. We take this for granted because, even though we know better, the assumption is somehow built into us that the new homes will not be needed for very long. Settlements of still newer houses will be built further along the trail. And we move always through a landscape which we secretly feel we can live without. It is already so far gone in polluted ugliness that is can hardly matter if it is made still uglier. Americans have gotten used to leaving their garbage wherever they happen to drop it. They
cannot take pride in an environment which looks to them like enemy territory. And it is part of our national heritage to suppose that they will never stay in one place long enough to be obliged to police the area and bury our beer cans.

The experience of driving by car from coast to coast is a case in point. The very physical structure of the drive shows that we regard the environment we are passing through as not worth looking at, and as foreign to us as the wilderness must have seemed to the first pioneers. The whole requirement imposed by the superhighway system is that we drive as fast as we dare and for as long as we can stand the strain to get where we are going as quickly as possible.

In any case, it is no wonder that each new generation of Americans appears to be more insensitive to the environment than the last. We seem able to survive within it only by making fortresses of our homes and staying inside them as much as possible. We are the most house-bound and house-proud of nations because there is so little worth leaving the house for. We suffer from having no village life or pub life or cafe life or market-square life. It is because we lack these things that our social life is so mechanical and lacking in warmth and ease. People do not just happen to come together or decide on impulse to drop in on one another. They are imported into each other's houses after the host and hostess have formally placed an order for them well in advance of the desired date of delivery. When
they arrive, the whole effort of hospitality is to get them drunk as quickly as possible so that they will soon not notice or care that they have nothing to say to one another. A vital social life depends upon the ability of a people to move freely and familiarly through their physical environment, to feel at home in it. But because we see our environment as foreign and ugly, we feel removed both from ourselves and from other people. For the psychological habit of holding one's self apart from unfriendly surroundings becomes a habit of social relations.

Perhaps it is true that there has really never been very much to do in this country except work. During the period of colonization, idleness may have been a threat to the developing economy and a sin against God. But the Protestant work ethic is the outgrowth of more than material need and religious belief. We have worked in order to distract ourselves from the fact that there are so few resources here for productive and satisfying leisure. We have created a whole society which is apparently able to go on forever distracting itself with objects and inventions.

1. According to this author, when and where has the civilized life existed in this country?

2. According to this selection, the hard work and cleverness formerly used to push back the physical frontier is now used how?
3. "It is already so far gone in polluted ugliness that it can hardly matter if it is made still uglier." This quote is an example of which of the following?
   a. Factual statement that can be proven or disproven  
   b. An opinion or interpretation

4. "By the time most immigrants came to the country, the ideals of political and religious freedom had been largely replaced by a promise of unlimited opportunity to improve the material conditions of life." This quote is an example of which of the following?
   a. Factual statement that can be proven or disproven  
   b. An opinion or interpretation

5. "But then, right at that point, something in our evolutionary mechanism went very wrong, and has remained very wrong ever since." This quote is an example of which of the following?
   a. Factual statement that can be proven or disproven  
   b. An opinion or interpretation

THE AMERICAN CHARACTER

Henry S. Commager

Let us consider some common denominators in the American character.

Perhaps the most common, and the most abundant, is carelessness. The American is careless in his manners, his dress and his address. He is careless about his house and garden. And he is careless in his social relationships, shuttling off old and taking on new with utmost casualness.

The American is, on the whole, openhanded, generous and hospitable. No other people pour so much money into churches,
schools, hospitals and other charities. No other has given so freely to help less fortunate people around the globe. While it is true that a system of tax exemptions makes it easy for Americans to be generous, it is suggestive that American tax laws are designed to encourage giving.

For two centuries visitors from the Old World have paid tribute to American hospitality. The American - as Denis Brogan has observed - was the first to make the term "stranger" a word of welcome. Along with material generosity went the quality of not holding a grudge. It is not without importance that America achieved nationalism without turning to national enmity. Americans did not nurse a grudge against Britain in the nineteenth century, nor have they been hostile toward Germany or Japan in the twentieth.

Though they have won most of their wars, they have never imposed a vengeful peace upon the defeated. Southerners fancy that the North imposed a "Punic peace" upon them in 1865. But the fact is that when the war was over, no one lost his life because of the rebellion and Southerners were back in Congress within a year after the last battle was fought. Compare this with what happened to unsuccessful rebels in Scotland, Ireland, France, Italy, Spain, Cuba or Russia in modern times.

Much of American generosity springs from good fortune and abundance. These account too for a third American trait: self-indulgence. The American dearly loves comfort and is acquiring
a taste for luxury. He pampers himself, his wife and his children. He rides instead of walking; overheats his house, his office and even his car. He thinks himself entitled to frequent and long vacations - to summers in the North Woods and winters in Florida. He spends enough on tobacco, liquors and cosmetics to pay off the national debt. What is achieved by the cigarette, the vacation in Florida, the electric mixer and the new car is happiness.

One of the more likeable American traits is gregariousness. It is doubtless a product of a frontier environment where men banded together to conquer nature and isolation. Americans like to do things together and they take comfort in numbers. They are the world's most enthusiastic joiners. The American tends to distrust the man who lives to himself, who prefers reading to conversation on a bus or plane, and who does not join fraternities or clubs.

On one trait almost all European critics are agreed: materialism. Is the American in fact more materialistic than his European cousins? Perhaps he is merely more honest. He is intensely aware of the material world in which he lives, the world of nature abundance, the world of industry and business. He is conscious too of size and space. Nor should it be forgotten that the process of becoming American was in part one of identification with the natural environment.

History, too, emphasizes material considerations. If the new nation was to preserve its independence and realize its demo-
ocratic potentials, it had to grow and grow fast. No wonder the American has always been fascinated by size and figures. No wonder he takes pride in the largest lakes and the tallest trees. It is fair to add that fascination with material growth has not in fact made the American more materialistic than the Frenchman or the German.

The American is confident and self-confident. He is a yes-sayer to life. Yet self-confidence melts easily into self-satisfaction and this into self-importance, and the American has rather more than his share of both. He believes his country has reached the heights of civilizations and that the "American way of life" is a moral rather than merely a social condition. The American constantly says he enjoys the highest standard of living, though the Scandinavian countries and Canada have a higher one. He says his medicine is the best in the world, though his country is eleventh in infant mortalities. He thinks his form of government is without doubt the best, though not one of the some fifty new nations formed since 1950 has seen fit to copy it.

This attitude expresses a deeper trait: the habit of supposing that the Americans are better-than-average people, that they are somehow free from the limitations of history. The American congratulates himself on his belief in political and social equality even while excluding one-tenth of the people from the benefits of equality. If there are embarrassing chapters in his
history, he ignores them. Other nations may want to expand their
control, but not the American. Ousting the Indian, the Spaniard,
and the Mexican somehow doesn't count. Other nations are mili-
taristic but not the United States, even though it engaged in
half a dozen major wars and, during the nineteenth century, almost
continuous Indian wars.

6. According to Commager, what does the American have "more
than his share" of?

7. According to Commager, what accounts for American self-
indulgence?

8. "... and Southerners were back in Congress within a year
after the last battle was fought." This quote is an example
of which of the following?
   a. Factual statement which can be proven or disproven
   b. An opinion or interpretation

9. "... not one of the some fifty new nations formed since
1950 has seen fit to copy (the American form of government)."
This quote is an example of which of the following?
   a. Factual statement that can be proven or disproven
   b. An opinion or interpretation

10. "One of the more likable of American traits is gregarious-
ness." This quote is an example of which of the following?
   a. Factual statement that can be proven or disproven
   b. An opinion or interpretation

11. "(America) is eleventh in infant mortalities." This quote
is an example of which of the following?
   a. Factual statement that can be proven or disproven
   b. An opinion or interpretation

12. "(Many Americans, as a group) spend enough on tobacco,
liquors and cosmetics to pay off the national debt." This
quote is an example of which of the following?
a. Factual statement which can be proven or disproven
b. An opinion or interpretation

13. "...the process of becoming American was in part one of identification with the natural environment." This quote is an example of which of the following?

a. Factual statement that can be proven or disproven
b. An opinion or interpretation

14. A flaw in some articles is the tendency to overgeneralize. Does Aldridge overgeneralize?

a. Yes
b. No

If you answered "Yes", give two examples from his selection to support your answer.

15. Does Commager overgeneralize?

a. Yes
b. No

If you answered "Yes", give two examples from his selection to support your answer.

16. Listed below are several topics that might be discussed in relation to Americans and their behavior. Circle the letter before each topic that both authors discuss in their analyses of American Character.

a. Materialism (item 37)
b. Selfishness
c. Spirituality
d. Over-confidence
e. Social relation with fellow Americans (item 38)
f. Luxury (item 39)
g. Identification with environment (item 40)
h. Generosity
i. Attitudes toward the home (item 41)
17. Discuss as concisely as possible the similarities and/or differences in the authors' views on each of the topics you selected in Question #16. Be sure the letter accompanying that topic also accompanies your discussion of their views on that topic.

18. Based upon these two selections, what conclusion can the reader come to concerning American Character?

19. What are your reasons for reaching this conclusion?