

THE DEVELOPMENT AND DECLINE OF  
CRITICAL THINKING ABILITY

---

A Thesis  
Presented to  
the Faculty of Graduate Studies and Research  
University of Manitoba

---

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Arts

---

by  
Celia Madeline Friend  
April 1958

## ABSTRACT OF THESIS

A review of the literature has revealed that experimental evidence of the role of age in higher thought processes is extremely meagre. In particular, the process of critical thinking has been severely neglected although its importance in our educational system and in our modern way of life has received a great deal of emphasis.

The present investigation deals with the nature of the development and decline of the critical thinking ability and seeks to study specifically the age at which the peak of critical thinking ability occurs, when it begins to decline and how pronounced is this decline. The course of development and decline of the various subabilities making up critical thinking is also studied, and consideration is given to the pattern of this development and decline for each of the different subabilities.

The Watson-Glaser Critical Thinking Appraisal Test was administered to a group of 484 subjects ranging in age from 12 to 89 years and representing diverse educational, occupational and economic backgrounds. This test measures the ability to draw inferences, to recognize assumptions, to use deductive thinking, to make interpretations of data and to evaluate arguments.

It was found that critical thinking ability shows a

progressive increase from late childhood through to the mid-twenties, holds up to 35 years and then declines progressively through to the seventies. The five subabilities were found to show a differential pattern of development and decline.

Although the older subjects, as a whole, scored poorly in critical thinking, many of them retained a high level of performance, which was as high or higher than that of young adult subjects. The poor performance in critical thinking in the elderly subjects is due, in part, to the two factors of rigidity in thinking, and low objectivity. In answering the test problems, older subjects tended to choose the absolutes of true and false, rather than to consider intermediate degrees of truth and falsity or other alternatives. The older subjects also allowed emotional feelings and biases to reduce the objectivity of their thinking about certain problems.

## TABLE OF CONTENTS

CHAPTER	PAGE
I. THE PROBLEM AND INTRODUCTION . . . . .	1
Statement of the Problem . . . . .	1
Introduction . . . . .	2
Nature of Critical Thinking . . . . .	3
Historical Background . . . . .	7
Studies Concerned with Higher Thought	
Processes of Children . . . . .	8
Appearance of Reasoning . . . . .	8
Development of Reasoning Processes	
in Children . . . . .	10
Studies of Higher Thought Processes	
in Adolescence . . . . .	13
Studies of Higher Thought Processes	
in Maturity and in Old Age . . . . .	18
Critique of Previous Investigations . . . . .	22
II. THE INVESTIGATION: MATERIALS, SUBJECTS	
AND PROCEDURE . . . . .	24
The Problem . . . . .	24
The Test: The Watson-Glaser Critical	
Thinking Appraisal . . . . .	24
Nature and Purpose of the Test . . . . .	24
Reliability and Validity . . . . .	27

	iii
CHAPTER	PAGE
Subjects . . . . .	28
Procedure . . . . .	33
Administration of the Test . . . . .	33
Scoring of the Test . . . . .	34
III. THE RESULTS AND DISCUSSION OF RESULTS . . . . .	36
Results of the Investigation . . . . .	36
Age Changes in Critical Thinking and its Subabilities . . . . .	36
Ability to Draw Inferences . . . . .	40
Ability to Detect Assumptions . . . . .	40
Deductive Thinking . . . . .	40
Ability to Interpret Data . . . . .	40
Ability to Evaluate Arguments . . . . .	41
Individual Differences . . . . .	41
Rigidity vs. Flexibility in Critical Thinking . . . . .	42
Objectivity in Critical Thinking . . . . .	44
Discussion of Results . . . . .	48
IV. SUMMARY AND CONCLUSIONS . . . . .	56
APPENDIX "A". The Watson-Glaser Critical Thinking Appraisal Test, Form Bm . . . . .	60
APPENDIX "B". Answer Sheet for the Watson-Glaser Critical Thinking Appraisal Test . . . . .	61

CHAPTER	PAGE
APPENDIX "C". Raw Scores Obtained by 484 Subjects in a Test of Critical Thinking Abilities . . . . .	62
APPENDIX "D". Occupations of 484 Subjects in a Test of Critical Thinking Abilities . . . . .	64
APPENDIX "E". "Emotionally toned" and "Neutral" Items of the Watson-Glaser Critical Thinking Appraisal Test . . . . .	66
BIBLIOGRAPHY . . . . .	67

## LIST OF TABLES

TABLE	PAGE
I. Distribution of 484 Subjects in a Test for Critical Thinking According to Sex, Median Age and Years of Education . . . . .	29
II. Occupations of 484 Subjects in a Test of Critical Thinking Ability . . . . .	30
III. Age Changes in Critical Thinking (Composite) and in its Five Subabilities, Mean T-scores . . . . .	37
IV. Percentage of Subjects in Different Decades Obtaining Scores Equal to or Above the Peak Mean of 70.34 Obtained by the 20 - 39 Age Group . . . . .	42
V. Mean Number of Choices by Three Age Groups of Five Possible Answers to Questions Requiring Inference . . . . .	44
VI. Objectivity Indices for Three Age Groups, Under 25 Years, 35-45 Years, and Over 60 Years . . . . .	47
VII. "Emotionally toned" and "Neutral" Items of the Watson-Glaser Critical Thinking Appraisal Test . . . . .	66

## LIST OF FIGURES

FIGURE	PAGE
1. Age Changes in Critical Thinking Ability (Composite of Five Subabilities.) . . . . .	38
2. Age Changes in Ability to Draw Inferences and to Detect Assumptions . . . . .	38
3. Age Changes in Deductive Thinking and in Ability to Interpret Data and Weigh Evidence . . . . .	39
4. Age Changes in Ability to Evaluate Arguments . . . . .	39

## CHAPTER I

### THE PROBLEM AND INTRODUCTION

#### I. STATEMENT OF THE PROBLEM

In the field of psychology, the area of human thought and judgment has been studied quite extensively but it appears that the role of age has been seriously neglected. Various investigators have interested themselves in the origin of higher thought processes in early childhood and some study has been made of the development of these thought processes in later childhood and adolescence. Past these periods, however, the literature concerned with age and higher thought processes has been meagre indeed.

On the whole, very little experimental study has been made of such a higher thought process as critical thinking, except as related to general intelligence or specific capacities. The research that has been done specifically on critical thinking ability is largely restricted to adolescents.

It is the purpose of this study to investigate the nature of the development and decline of critical thinking ability throughout the life span. More specifically, the study hopes to answer the following questions: At what age

does the peak of critical thinking ability occur? When does this ability begin to decline and how pronounced is this decline? What is the course of development and decline of the various subabilities making up critical thinking? Is the pattern of development and decline the same, or different for the different subabilities?

## II. INTRODUCTION

The necessity for the present study was felt after a review of the literature revealed a dearth of experimental evidence for the numerous speculations and generalizations regarding the relationship between the higher thought processes and aging. In particular, the study of critical thinking appears to have been severely neglected while, at the same time, a great deal of emphasis has been placed upon its importance in our modern way of life.

Our present complex society demands a high degree of social understanding and critical-mindedness which is necessary to make intelligent judgments about public issues and private interpersonal relationships. Life in a free-enterprise state with its emphasis upon economic factors demands constant attention to advertising and propaganda of all sorts. Complexities of society demand constant weighing of values on the part of the individual. Existence within a democracy means confronting controversial issues and making wise choices

after varying points of view have been criticised and evaluated. It would seem, then, that critical thinking is one of the first requirements of a citizen in a democracy.

### Nature of Critical Thinking

Critical thinking is one type of higher thought process, others being such processes as perceptual, associative and creative thinking, as well as concept formation. Various definitions of critical thinking are possible. Most simply, it can be considered as the use of logical reasoning and the avoidance of common fallacies in judgment. Amongst the different types of higher thought processes, critical thinking can best be placed between problem-solving and creative thinking, as it is closely related to both. It is one of the activities considered essential to problem-solving at the hypothetical stage, and it occurs toward the end of creative processes, when appraisal is employed.

Not all problem-solving or creative thinking involves the processes of critical thinking, of course, but where evaluation, comparison to some standard and the examination of ideas and products are called for, critical thinking must be employed. In relation to creative thinking, it complements spontaneous, free-flowing, uninhibited thinking and the exercise of constructive imagination. The creative thinker needs to review his creations with a judiciously critical eye in order to improve or discard, as needs be. This involves

the critical thinking abilities of discerning observation and insightful interpretation of minimal cues, with the avoidance of "shrewd guesses" which at best are only hypotheses, subject to verification.

Russell (65) suggests that critical thinking ordinarily involves four conditions: (1) a knowledge of the field in which the thinking is being done, (2) a general attitude of questioning and suspended judgment; a habit of examining before accepting, (3) some application of the methods of logical analysis or scientific inquiry, and (4) taking action in the light of this analysis or reasoning.

Other investigators have suggested a wide range of abilities which are to be considered as components of critical thinking: introspection, comparison, appraisal and even such affective factors as a measure of belief (1, 3, 5, 20, 58, 70). Some, such as DeBoer (15) emphasise the establishment of criteria for judging as being an essential factor in the study of this particular type of thinking process. Still others (22, 60) have concerned themselves with setting forth steps for persons doing critical thinking about printed matter or problematical situations. These various studies have led Russell (66) to define critical thinking as "the process of examining both concrete and verbal materials in the light of related objective evidence, comparing the object or

statement with some norm or standard, and concluding or acting upon the judgment made."

A number of investigations have revealed that critical thinking is correlated to some degree with intelligence. However, all of the correlations are quite low, ranging from .00 to .46 (10, 20, 27, 30).

Watson and Glaser (75) maintain that this somewhat low correlation of critical thinking and general intelligence is due mainly to the non-abstract quality of critical thinking. They point out that basic intellectual abilities are measured, in the main, by abstract tests, but that critical thinking ability can best be measured by providing problems and situations of realistic types, similar to those which a person might encounter in daily life as he works, reads the newspaper, hears speeches and participates in discussions.

Undoubtedly the most extensive and best work being carried on at present on critical thinking is being done by Watson and Glaser. They maintain that the ability to think critically involves three things: an attitude of desiring supporting evidence before accepting opinions and conclusions as true, a certain degree of knowledge of logical methods by which different kinds of evidence can be weighed and which help to arrive at warranted conclusions, and skill in employing such attitudes and knowledge. Instead of jumping prematurely to conclusions, a critical thinker examines beliefs

or proposals in the light of supporting evidence and relevant facts. Critical thinking requires that he be able to comprehend and use language with accuracy, clarity and discrimination. It also requires ability to recognize problems in the first instance and to find workable means towards their solutions, to gather and arrange pertinent data, to appraise evidence and to evaluate arguments. In recognizing the existence (or non-existence) of logical relationships, the critical thinker is enabled to draw warranted conclusions and make proper generalizations. He needs to reconstruct patterns of beliefs on the basis of wider experience and to render accurate judgments about specific things and qualities in everyday life.

According to Watson and Glaser, the "tools" which a critical thinker employs are those of logical reasoning, generalization, deduction, induction, inference, the recognition of assumptions, discrimination of arguments and interpretation of basic data. Working from these views, Watson and Glaser have devised a test (The Critical Thinking Appraisal Test) which, from all accounts, seems to be an excellent measure of the critical thinking process and its various facets, or subabilities. This test was used in the present investigation to study the course of development and decline of critical thinking abilities.

The thesis begins with a discussion of the historical

background of the problem, presenting some of the theoretical work and relevant research. Following this introductory discussion, the procedure is described, and a description of the Watson-Glaser Critical Thinking Appraisal test is given. The results of the investigation are then presented, together with a discussion of their implications. The concluding section deals with a summary of the results and conclusions drawn from the investigation.

### III. HISTORICAL BACKGROUND

Interest in the higher thought processes has been expressed, for the most part, speculatively. Literary comments upon these intellectual processes have been made for centuries. The early Greeks were, of course, greatly interested in processes of logical reasoning and in most of the factors involved in critical thinking which were heretofore mentioned. They did not, however, interest themselves in the developmental stages of these processes within individual persons. In so far as comment upon critical thinking is concerned with the aging process, the emphasis has been upon the growing conservatism and intellectual rigidity of the aged. Since mental abilities have most often been regarded as dependent to a marked degree upon the physical condition of the individual, it is not surprising that their decline has been considered parallel to that of physical decline with

age. Much of the literature is concerned with general declining trends of reasoning abilities after the age of 60 and, in particular, with a marked drop at 65. There is little, if any, supporting evidence for the arbitrary choice of these ages, however. Likewise, speculation as to the beginning of critical thinking and the peaks of its efficiency are, in the main, unsupported by experimental studies.

#### Studies concerned with Higher Thought Processes of Children

Practically all experimental investigations of reasoning, judgment and component parts of critical thinking have been carried out with children and students. These have centred mainly around such questions as: At what age does the ability to reason first appear? How does reasoning ability develop as the child becomes older? Is there any essential difference between the reasoning ability of children and that of adults? How may critical thinking abilities best be developed through an educational programme?

Appearance of reasoning. Probably the greatest amount of research on children's thinking has been expended upon the establishment of an age level for the beginnings of reasoning processes. There has been little concurrence in this regard, some investigators claiming evidence of reasoning as early as two years, others denying its existence before the age of seven years.

Heidbreder (38), for instance, considers that some reasoning ability is evident at the age of two and one-half years. In a problem-solving experiment using 10 subjects of three years, 10 of four years, 10 of ages six to ten, and 10 adults, she found that none of the youngest group actually gave a reason for his solution and only five per cent of the four-year-olds had reasons, but of the older two groups, all subjects presented evidence of reasoning ability. Roberts (62) claimed the appearance of reasoning ability occurred at a little over three years of age. He tested 45 subjects from two to five, as to their ability to discover and apply to new situations the solving principle of a given problem. All subjects were successful in achieving a solution and the youngest who could give a verbal generalization was three years, four months old. Frequency of verbal generalization increased with chronological and mental age, Roberts reported, and the ability to apply the principle to new situations was evidenced by all but the two-year-olds.

From his studies of sequences in conceptual thinking, Griffiths (32) claims that the beginning of reasoning is between three and six and one-half years. Similar conclusions are reported by Goodenough (31) and Russell (64). Conceptual thinking was also the concern of Hurlock (40) who, while setting no arbitrary age level, noted that specificity of concepts increases with age.

The average age of the appearance of reasoning ability is claimed to be a little later than this, by Harter (34) who reports that this ability appears at about 60 months. In a problem-solving experiment involving 54 children from three to six years, she tested their ability to solve simple problems with or without overt trial-and-error. Here it was found that the correlation between chronological age and success in solving problems was  $.72 \pm .06$ , and for mental age the correlation was  $.57 \pm .08$ . Implicit solutions (reasoning), she reported, became more frequent with increase in age.

Finally, Piaget (57) claims that there is no evidence of reasoning processes in children younger than seven years. He conceded that children can reason at a much earlier age than twelve, however, if the reasoning was defined in terms of verbal problem-solving at a given level of experience and development.

Development of reasoning processes in children. Having surveyed the literature on the age of appearance of reasoning and other higher thought processes, we will now turn to the literature concerning their development with age. The early investigations of Burt (12), for instance, were concerned with reasoning and conceptual thinking in children and dealt with what he considered the essential organization and hierarchy of thinking processes. He used factor analysis to establish a series of these abilities. By 1923 he had done

considerable work on the age levels of reasoning ability and compiled a number of related tests. In 1919 he reported that children, as they grew older, could solve problems with an increasing variety of subject matter and that the subject matter seemed more important to the children than the logical form of the questions. He concluded that "the development of reasoning appears to consist essentially in an increase in the number, variety, originality and compactness of the relations which his [the child's] mind can perceive and integrate into a coherent whole" (11).

Piaget (57) has done a considerable amount of research on children's thinking processes. His work, however, is largely of an observational nature, involving comparisons of children's thinking with that of adults, or records of children's verbalized concepts. Spontaneous speech exhibiting conceptual thinking, he claimed, was egocentric for the first seven or eight years, became more analytical between seven and eleven and did not evidence syllogistic thinking until after the age of twelve. Deutch (17) agreed with this general observation and felt that children use logic that is within their intellectual grasp. Adequacy of this type of thinking, he said, is related not to the age factor alone but also to sex and socio-economic status. Hazlitt (35) also concluded that in Piaget's experiment on judgment and reasoning in the child, age was not the differentiating factor. Rather, she

says, it is the information held by the child. Piaget, she claims, overvaluated verbal expression as a measure of thinking.

Moore's 1929 study of reasoning of the syllogistic type (53) led him to conclude that development of logical ability is a function of chronological age. Using 205 children from six to twelve years of age, and 105 college freshmen, he tested them with regard to verbal reasoning and ability to discern autistic and logical fallacies and found steady improvement in each case, at every successive age level. Although no child under eight years could discern autistic fallacies, all of his subjects achieved some degree of reasoning and logical ability.

Maier (48) tested the ability of children to combine or integrate two isolated experiences. Thirty-nine children from 43 to 95 months were given a maze test, from which Maier found a steady increase with age in the average accuracy and the percentage of subjects meeting the criterion of five correct trials out of ten. He felt it unwise to analyse the results in detail, or to plot curves of development, but felt that the results did show that goal achievement, as a result of combining two isolated experiences, is rather late in maturing. He found that it is rarely developed to any marked extent under the age of six years.

A similar increment in the effective use of reasoning with an increase in age was found by Gellerman (29) who tested 38 children from three to thirteen years old, and 25 college students in a double alternation problem. No pre-test information was given to the subjects regarding the nature of the problem. Gellerman reports that beyond the age of four years there was at successive age levels a gradual decrease in the number of trials required for learning the maze.

Werner and Kaplan (78) used 125 children, from eight to thirteen years of age in an investigation of concept formation and reported that remarks of critical evaluation increased sharply between the ages of ten and eleven, with older children modifying their solutions more readily after self-criticism.

These and other similar genetic approaches to the questions regarding children's thinking and reasoning processes suggest that they first appear sometime between two and seven years, and develop quite rapidly thereafter.

#### Studies of Higher Thought Processes in Adolescence

The literature on higher thought emphasises the great importance of the development and effective use of such processes as concept formation, verbal and abstract reasoning, inductive and deductive thinking, associative, creative and

critical thinking at the high school and college levels. Most writers have stressed the importance of these as objectives of education. Therefore, many of the reported investigations are concerned with the improvement of these thinking processes in the course of teaching specific subjects in such areas as reading (15, 23, 28, 44, 45, 46, 73), science and mathematics (20, 24, 41, 72, 79) and social studies (9, 19, 52, 63).

Several studies of specific higher thought processes have been carried out with older children and adolescents on such topics as inductive reasoning (56, 74), abstract and concrete reasoning (16), verbal reasoning, and problem-solving (18). For the most part, however, the age factor has been ignored in these studies.

The growth of insight at this age has been studied by a number of investigators. Pyle's experimental study (59) of the development of insight in adolescence, as measured by a total of five poems, led him to report a steady upward trend in the effective use of this ability, from Grades III to VIII. Adolescent ability to obtain insights was also investigated by Franklin in 1928 (26). He gave a series of objective tests to 637 children and adolescents and reported gradual gains in insight between the mental ages of eight and eleven. There followed large increases up to about

fourteen years, after which the increase was not so noticeable and the rate of increase fluctuated somewhat. He reported also that there was an observable increase in the ability to understand allegories and double meanings at this age group.

A somewhat similar experiment to assess the ability to obtain insights was undertaken by Shaffer in 1930 (67). As in the Franklin study, the subjects were arranged by mental age rather than by chronological age. This study, involving the power to "see through" situations depicted in cartoons, revealed that there was very little increase in this ability from the mental age of nine to eleven, but a steady and steep climb from eleven to sixteen, after which there was a tendency for the increase to level off again. He concluded that it was in this ability to interpret meanings that the adolescent is most distinctively different from the child.

To trace the growth of ability to match principles with illustrations and to evaluate explanations, Edwards (20) conducted an investigation among students from Grade IX through the second year of college (a span of six years from approximately fifteen to twenty years of age.) He found that the average gain in the ability to evaluate explanations was only from 16.8 to 23.5 points, out of a total possible score of 36. In this ability, however, the gain was fairly

commensurate with age increase. For the ability to match principles with illustrations, however, there was some dropping off of scores at the approximate age of eighteen or nineteen years (first year of college.)

Inductive thinking of adolescents was the particular interest of Shukla (68) who employed 357 subjects, thirteen to fifteen years of age, in verbal "reasoning" tests and found that all of them used inductive reasoning but, at this age, were unable to tell how they made their discriminations.

An age scale to measure successive steps of the higher thought processes has been suggested as a desirable object of research by Cole (14), who points out that the Binet "comprehension" items measure successive steps of reasoning, judging and obtaining insights and, through a scatter of eight age levels (from age three and one-half to Superior Adult) can be used as a measure of ability to judge and to draw inferences. She also suggests that an age scale could be made by using the kinds of jokes that seem amusing to children of different ages, and cites the Binet scale of "absurdity" items which could be the basis for such a measure.

From such studies as that of Bradford (7) who investigated the critical attitudes of twelve and thirteen-year-olds as early as 1925, it appears quite evident that distinctive features of this type of thinking are already formed and in use. Bradford reported that 57 per cent of his subjects

supplied answers even to unsolvable problems and over 75 per cent of them gave answers to at least some of the problems of all types which were set before them. At this age level, he concluded, students tended to conform to habit and to adult authority rather than to think independently. Harris (33), by analysis of verbatim reports, concluded that twelve and thirteen year old pupils can think critically about classroom procedures which are of direct concern to them, but not about such topics as international trade.

The only study specifically concerned with critical thinking ability was carried out by Glaser in 1941 (30). Glaser devised the Critical Thinking Appraisal Test which he administered to a total of 5,969 subjects (2,856 Grade XI students with a mean age of 16.9 years, and 3,113 Grade XII students with a mean age of 17.8.) He found that the mean critical thinking score (as measured by his Critical Thinking Appraisal) of the two groups was almost the same. Apart from this investigation, the evidence of change in critical thinking abilities in the late adolescent age group is unsupported by any experimental findings. It is the seventh revised edition of this Critical Thinking Appraisal Test which is used in the present investigation.

Studies of Higher Thought Processes in Maturity and Old Age

Experimental findings with regard to further development of the higher thought processes in adulthood, and particular investigations into the points of peak and decline of these abilities are meagre indeed. However, some studies have been undertaken, which cover the general field of mental development and these yield some indications of the mature adult's use of such abilities as those involved in deductive thinking, comparing and making judgments, and critical thinking.

An investigation covering the whole range of mental development from infancy to maturity and from low grade mental defectives to adults of superior mental capacity was the particular concern of Foulds (25) who studied 5,000 skilled and unskilled males in a test of Progressive Matrices (Raven, 61). This test consists of a number of problems requiring solution by comparison, reasoning by analogy and logical deductions. His findings led him to conclude that a man normally possesses his greatest knowledge and clarity of thinking between the late 20's and early 40's. Before this, his capacity for clear thinking is at its maximum, but he lacks the breadth of knowledge necessary to put his ideas into practice effectively. By the early 40's, his knowledge is usually reaching maturity but his capacity to understand fresh methods of thinking has begun to decline.

In addition, it was found that he took longer to grasp new methods of working and, in coping with new problems, relied more and more upon existing knowledge.

The use of deductive thinking was the object of Speakman's study in 1952 (69). This investigation, using 67 subjects ranging in age from 20 years to 87 years, yielded some interesting comparisons between older and younger subjects. The study was of the ability to relearn the values of postage stamps which had been issued in changed colours. One of the sub-tests involved the use of deduction. Here, a clue card was issued to each subject and by its use he could deduce the value of the stamps, by a mathematical process of elimination. The use of the clue card was not actually necessary, however, nor was it obligatory. Speakman reported that the proportion of people using the clue card for aid in deductive processes tended to diminish slightly after 40 and more abruptly after 60 years. The number using logic after this age was found to be negligible. Efficiency of deduction among those who did use the clue card tended to rise just after 40, but fell thereafter. In all age groups, the percentage of actual improvement over possible improvement was higher for the subjects who used logic than for those who did not. Whereas the younger people seemed to regard the clue card as a useful aid, finding logical analysis easier than

trying to create order among a mass of partly-remembered meanings, many of those over 60 quite obviously regarded it merely as another complication in an already difficult problem. Many older subjects tended to make judgments that "felt right," and either tried to twist data to fit, or ignored data which did not fit. Many seemed to lack the ability to carry on significantly connected trains of thought of the type necessary to solve the problem.

Decline in the use of logical methods found in the Speakman study parallels that of Allan (76) who, using two groups of subjects (35 years and under, and 36 years and over,) presented controversial statements of political and moral issues. Some of these were incompatible and the subject was asked to eliminate these, explaining their incompatibility. Allan found a tendency among the older group to comment on the material from their experience instead of making deductions from the material itself. He concluded, however, that it was not possible to say how far this apparent inability of older subjects to use logical analysis would be general to all tasks, and decided that it well may be that, in actual life situations, logical deduction becomes a habit linked to certain familiar situations where its utility is proven. There is a tendency not to apply logic to unrelated or novel contingencies, or to political or other material where sound premises are usually lacking

or emotionally unacceptable.

Age changes in logical thinking has been studied by Welford (77). His subjects were required to detect fallacies and inconsistencies in a number of statements, and to draw deductions. He found that "older subjects tended not to draw logical deductions based strictly on the statements as given, but to introduce supplementary premises, or to confine themselves to comments upon the statements." He maintained that this failure to make deductions is probably due to faulty organization and interpretation of data by the central neural mechanisms.

Critical thinking calls for a certain degree of flexibility in thinking. A test measuring this flexibility was devised by Berg (4) in 1948. A group of young subjects of 17 to 28 years, and a group from 58 to 73 years were asked to sort cards into a number of piles. Sorting could have been done by several criteria, but the "correct" one was arbitrarily chose by and known only to the experimenter. When the subject chose it, this criterion was changed without his knowledge. The ability to detect these changes was the measure of flexibility. Berg found that not one of the older group could progress after the first shift, while the younger ones shifted readily.

The only study specifically concerned with the later developmental changes in critical thinking appears to be

that carried out by Watson and Glaser (75). Only ten subjects were used in this investigation. These were carefully selected trainees in Industrial Relations, all college graduates with advanced training and several years of work experience. Their ages ranged from 25 to 31 years (median 29.) Critical thinking scores for this special group ranged from 65 to 87, with a median of 79. This score is considerably higher than that obtained on the same test by high school students, and indicates that critical thinking scores can increase with age.

#### Critique of Previous Investigations

The preceding review of the literature on the higher thought processes has shown that numerous types of intellectual processes have been studied. For the most part, experimental investigations seem to have been confined to the establishment of age levels for the beginnings of such processes as reasoning in problem-solving situations, evidence of concept formation, the use of inductive and deductive thinking, logical analysis and the growth of insight. Studies have been almost entirely confined to children and students. Many of the findings have resulted from limited investigation of small groups or as subsidiary studies to other investigations. The role of age has been largely neglected, except for broad generalizations concerning

"general trends" of decline and the supposed parallelism of critical thinking abilities with those of general intelligence. Findings have, furthermore, been based mainly upon subjective appraisals or on observation of performance rather than measurement and statistical procedures. Conclusions have been drawn from comparisons between older and younger age groups, rather than from groups covering a wide span of years.

The available literature reveals that only one or two studies have dealt specifically with the critical thinking ability and these have largely ignored the age factor. Beyond an academic interest in the use of certain sub-abilities of critical thinking, and suggestions for the development of these abilities during school years, little has been done with regard to studying the age changes in critical thinking processes throughout the life span.

## CHAPTER II

### THE INVESTIGATION: MATERIALS, SUBJECTS AND PROCEDURE

#### I. THE PROBLEM

The previous discussion revealed the dearth of research into the question of the development and decline of critical thinking abilities. It was because of this paucity of data that the present investigation was designed. Its main purpose is to throw some light on the course of development and decline of critical thinking ability and of its more important subabilities.

#### II. THE TEST: THE WATSON-GLASER CRITICAL THINKING APPRAISAL TEST

##### Nature and Purpose of the Test

The Watson-Glaser Critical Thinking Appraisal (revised edition, 1952) was chosen for measuring critical thinking abilities in this investigation because it is the only good test available at the present time. It was designed to provide problems and situations requiring the employment of some of the important abilities involved in thinking critically. The total concept of critical thinking is appraised by measuring separately several of the important subabilities making up this intellectual activity.

Earlier editions of the Watson-Glaser Critical

Thinking Tests (1942) were offered in two batteries as a means for obtaining quantitatively, as well as qualitatively, a measure of the person's ability to think critically with regard to problems involving interpretation of data, drawing inferences, discriminating between strong and weak arguments, recognizing assumptions in reasoning, making warranted generalizations and applying logical reasoning. The seventh Revised Edition, used in this investigation, omits the latter two, as well as two paired tests regarding opinions, and presents the remaining five subtests in one battery so that completely objective scoring is possible.

The usefulness and significance of the items in this seventh or 1952 revision have resulted from successive refinements or replacement of items. This final revision arose out of various experimental studies with high school and college students and such specialized groups as graduates, personnel and administrative applicants, civil servants, business executives and industrial personnel. The edition used in this investigation is issued in two equated forms, Am and Bm, each containing 99 items and matched both in average difficulty and in the distribution of their difficulty coefficients. Form Bm was used for this study. A copy may be found in Appendix "A."

The five subtests, together with the abilities they measure are:

1. Inferences. (20 items.) This test aims to sample the person's ability to discriminate among degrees of truth or falsity or probability of inferences drawn from given data. It is a measurement of his inductive processes.

2. Recognition of Assumptions. (16 items.) This test was designed to sample the individual's ability to recognize unstated assumptions in given propositions or assertions. The score here is indicative of the degree of proneness to errors in reasoning or to fallacies. Well-founded assumptions are hypotheses, and the continuity of good reasoning depends upon recognizing the assumptions as such, followed by an estimation of the probability of truth of the assumption itself.

3. Deduction. (25 items.) This test was designed to measure the subject's ability to reason deductively from a given premise, to recognize the implications between propositions and to determine whether what seems an implication between one proposition and another is actually so.

4. Interpretation. (24 items.) This test was designed to sample the subject's ability to weigh evidence and to distinguish between unwarranted generalizations and probable inferences which, though not conclusive or necessary, are warranted beyond a reasonable doubt.

5. Evaluation of Arguments. (14 items.) This test

samples ability to distinguish between arguments which are strong and important to the question at issue, and those which are weak and unimportant or irrelevant.

### Reliability and Validity

The test as a whole has been shown to have adequate statistical reliability, this having been determined by both the split-half ( $r = 0.90$ ) and inter-form methods ( $r = 0.81$ .) The judgment of 35 persons, selected for their advanced training in logic and the use of language, was called upon to establish the key of correct answers to the test items. These judges agreed unanimously that the answers to the finally revised items are logically correct and that correct response to the items requires some of the most important abilities which are fundamental to critical thinking. The validity of the test and the key was corroborated by high correlation between success on individual items and on the total test, and in several tests against independent criteria. Watson and Glaser concluded that their test is an excellent measure of the ability to comprehend and reason logically about the kind of problems contained in the Critical Thinking Appraisal.

The Critical Thinking Appraisal is not an intelligence test as such. It differs from abstract tests which are designed to measure basic intellectual abilities. Watson and Glaser found, for instance, that the correlation between

their Critical Thinking Appraisal and the Terman-McNemar Intelligence Test was 0.70, when used in a study of 192 Grade XI students. In general, the correlations with intelligence tests tends to cluster around 0.45. Watson and Glaser point out that many persons who have superior mental ability as measured by an intelligence test, may make a relatively low score on the Critical Thinking Appraisal. But if a person makes a relatively high score on the Critical Thinking Appraisal, he is likely to score relatively high on an intelligence test.

### III. SUBJECTS

The main object in the choice of subjects was to obtain a representative cross-section of normal persons, ranging in age from late childhood through to senescence, and whose educational, occupational and economic status would be as widely diversified as possible. This diversification should also be represented within each age group, as well as throughout the entire population. Care was taken to keep a numerical balance with regard to a number of factors, the most important being:

1. Grouping. The number of subjects, ranging in age from 12 years to 89 years, totalled 484. These were divided into seven age groups, by decades, as shown in Table I below. At least fifty in each decade under 70 years

was deemed a minimum sample.

TABLE I  
DISTRIBUTION OF 484 SUBJECTS IN A TEST FOR  
CRITICAL THINKING, ACCORDING TO SEX, MEDIAN  
AGE, AND YEARS OF EDUCATION

Age Group	Total Number	Sex		Median Age		Years of Education (Mean)
		M	F	Yr.	Mo.	
Teens	63	28	35	16	- 11	10.5
20's	85	43	42	24	- 7	13.7
30's	95	48	47	35	- 0	10.9
40's	82	39	43	44	- 7	11.4
50's	67	31	36	55	- 3	11.4
60's	57	29	28	63	- 9	11.6 **
70 and over	35	11	24	74	- 5	9.7 *
	484	229	255			

\*\* : Only 28 specified their years of education.

\* : Only 10 specified their years of education.

2. Age. Selection was made so that the median age for each group would approximate the mid-decade mark (see Table I.) Accurate and honest reporting of age by the subject was relied upon. Wherever any doubt occurred in this respect, and if no check was possible, the test form was rejected.

3. Sex. An even balance in each group was the aim (see Table I.)

4. Education. Again, honest and accurate reporting was required of the subject. Years of formal education, including professional training, were counted, with a specified number of years allotted for various academic degrees. Table I shows that the subjects in the various decades have approximately the same educational background.

5. Occupation. Nineteen female and 45 male occupational categories, with additional subdivisions, were represented. The distribution, grouped into six main categories, is shown in Table II below. The complete list of occupations will be found in Appendix "D."

TABLE II  
OCCUPATIONS OF 484 SUBJECTS IN A TEST  
OF CRITICAL THINKING ABILITY

Age Group	Total Number	Occupational group					
		I	II	III	IV	V	VI
Teens	63	0	5	1	0	56	1
20's	85	3	40	3	1	19	19
30's	95	8	40	10	4	0	33
40's	82	6	43	4	6	1	22
50's	67	7	25	9	2	0	24
60's	57	8	24	4	3	0	18
70+	35	3	7	3	2	0	20

Group I: Professional. Higher executives, lawyers, clergymen, university professors, highly qualified scientists, company managers, dietitians, librarians.

- Group II: Semi-professional. Small business proprietors (druggists, grocers, jewellers, photographers,) clerical services (office clerks, accountants, bank tellers, stenographers, secretaries,) salesmen, insurance underwriters, junior technicians, laboratory assistants, nurses and teachers not in supervisory positions.
- Group III: Skilled Labour. Cabinet makers, carpenters, machinists, senior mechanics, foremen, police, carmen, linemen, railway engineers, farmers.
- Group IV: Unskilled Labour. Caretakers, construction helpers, domestics, waitresses, locker attendants, mechanical helpers, farm hands.
- Group V: Students, not presently employed.
- Group VI: Housewives, who did not specify particular employment outside of the home.

It will be noted that most categories are well represented throughout all the decades, with the obvious exception of students. Within each age group, a balanced percentage of occupations was the aim. Of the over-70 group, only ten were institutionalized. Eight of these were in a home for the aged and two were temporarily in a convalescent hospital.

6. Residence. The subjects were for the most part residents of the city of Winnipeg (347), although 83 tests were given in Saskatoon and 25 in Ottawa. Ten subjects were residents of other Canadian cities, 19 were from rural points in Manitoba, Saskatchewan and Ontario. It is impossible to state how many of the urban-domiciled were of rural or mixed geographical background, but undoubtedly many were

such.

No control was kept over intelligence factors, other than that only obviously "normal" persons were tested.

Tests were rejected for various reasons besides the previously mentioned doubt as to correct age. Except for one or two small classes of students, each subject was asked individually to complete the test form. The final 484 subjects represent about one-third of those asked. The remaining two-thirds refused on grounds of lack of confidence in their ability to cope with pencil-and-paper tests, disinterest, fear that the test was one of intelligence or that the "authorities" would learn their score. When assured of anonymity if they so desired, a number were later persuaded to take the test. Except for a few, each testee was supervised by the investigator or a trained assistant. The few exceptions were well-known and trusted persons who wished to complete the test form at their leisure and return it by mail or personal delivery. A number of others who wished to protect their anonymity from the group in which they were working, or from the investigator, were also allowed to return the form by mail, but a careful check was made of these and if any doubt arose as to the reported personal facts, or to the authenticity of the answer form being the work of the subject, it was rejected. There were very few of these.

## IV. PROCEDURE

Administration of the Test

For the most part, the test was administered to individuals or small groups of three or four persons. The instructions provided in the Critical Thinking Appraisal Manual (75) were followed, but given in somewhat more informal terms when used with individuals. The largest single group of testees numbered 38, a high school class chosen for its variations in economic and social background rather than for its homogeneity. Each member of this and other smaller groups of individually chosen subjects was supplied with a test booklet, separate answer sheet, and pencil, and instructed as follows:

"May I have your attention, please? Each of you has been given a test booklet and a separate answer sheet. Do not open the booklet or make any marks on it until I tell you to do so.

"This booklet contains five tests designed to find out how logically and analytically you can think.

"Each test is preceded by its own specific directions. When I tell you to begin, you will read the directions for the first test and study the sample questions until you know what you are to do. If you cannot readily determine what the directions mean, raise your hand and I will explain them to you. Do not ask questions about a test after you start to answer it. Help will not be given on any test you have already started to answer. Make absolutely no marks on the test booklet.

"For each question, decide what you think is the best answer. Then record your choice by making a black mark between the appropriate pair of dotted lines on the answer sheet. If you change your mind about an answer, be sure to erase the first mark completely. You may answer a question even when you are not perfectly sure that your answer is correct, but you should avoid wild guessing. Do not spend too much time on any one item. When you finish a page, go right on to the next one. Work as rapidly and as accurately as you can.

"In marking your answers, always be sure that the answer space is numbered the same as the question in the test booklet.

"Remember that you are to start reading the directions when I tell you to start. If you wish to change an answer, erase completely, and make no marks on the test booklet. After you have finished checking your answers, return the answer sheet to me. Are there any questions before we begin?

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

The testees were asked to write their names, occupation, education and exact age on the answer sheet. There was no time limit on the test. Time varied from about forty minutes to several hours. With some of the aged subjects who were incapacitated for such reasons as poor eyesight or arthritic hands, the investigator read the items and marked the choices for the subject.

### Scoring of the Test

Each answer sheet was scored by the use of the key provided with the test, and raw scores obtained in this manner, for each of the subtests as well as for the total test.

Since the maximum scores on each of the subtests varied, all raw scores were converted into Z-scores. These were then again converted into T-scores which represent the individual's score in a distribution with a mean of 50 and a  $\sigma$  of 10. All graphs are plotted in T-score units.

## CHAPTER III

### THE RESULTS AND DISCUSSION OF RESULTS

#### I. THE RESULTS OF THE INVESTIGATION

The raw scores obtained by the 484 subjects are given in Appendix "C." Their converted raw scores are summarized in graphical and tabular form in Figures 1-4 and Tables III-VI.

#### Age Changes in Critical Thinking and its Subabilities

Figure 1 and Table III indicate the course of development and decline of critical thinking ability (composite of five subabilities.) It can be seen that this ability is not too highly developed in teen-age subjects. However, it rapidly develops with the years, reaching a peak in the mid-twenties. Critical thinking ability then holds up well with age into the mid-thirties, after which it begins to decline steadily through to old age.

When consideration is given to each of the five subtests, there is little difference noted in the general patterns of peak and decline. Although some peaks are achieved later than at 25 years, none of the subabilities holds up later than the age of 35, as indicated in Figures 2, 3 and 4.

TABLE III  
 AGE CHANGES IN CRITICAL THINKING (COMPOSITE)  
 AND IN ITS FIVE SUBABILITIES--MEAN T-SCORES

Critical Ability	Teens	20's	30's	40's	50's	60's	70+
Inference	48.60	54.30	54.49	49.39	47.61	46.04	37.22
Recognition of Assumptions	46.66	53.64	53.83	51.07	49.08	45.60	42.57
Deductive Thinking	47.82	53.45	55.48	50.66	47.97	43.68	40.63
Interpre- tation	49.25	55.88	53.29	51.12	49.05	48.27	46.39
Evaluation of Arguments	47.42	54.91	53.83	50.32	49.73	44.80	43.13
Critical Thinking	47.95	54.38	54.18	50.51	48.69	45.68	41.99



Figure 1. Age changes in critical thinking ability (Composite of five subabilities.)

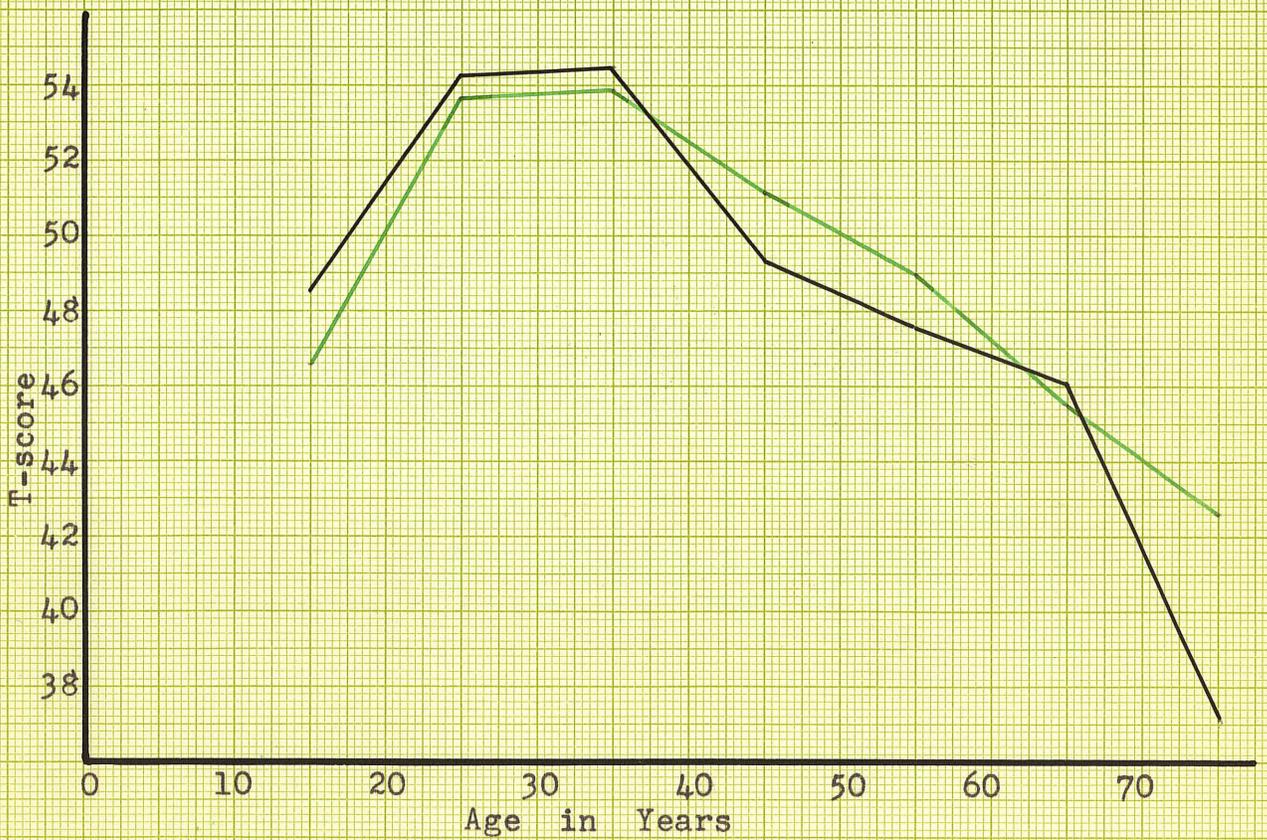


Figure 2. — Age changes in ability to draw inferences.  
 — Age changes in ability to detect assumptions.

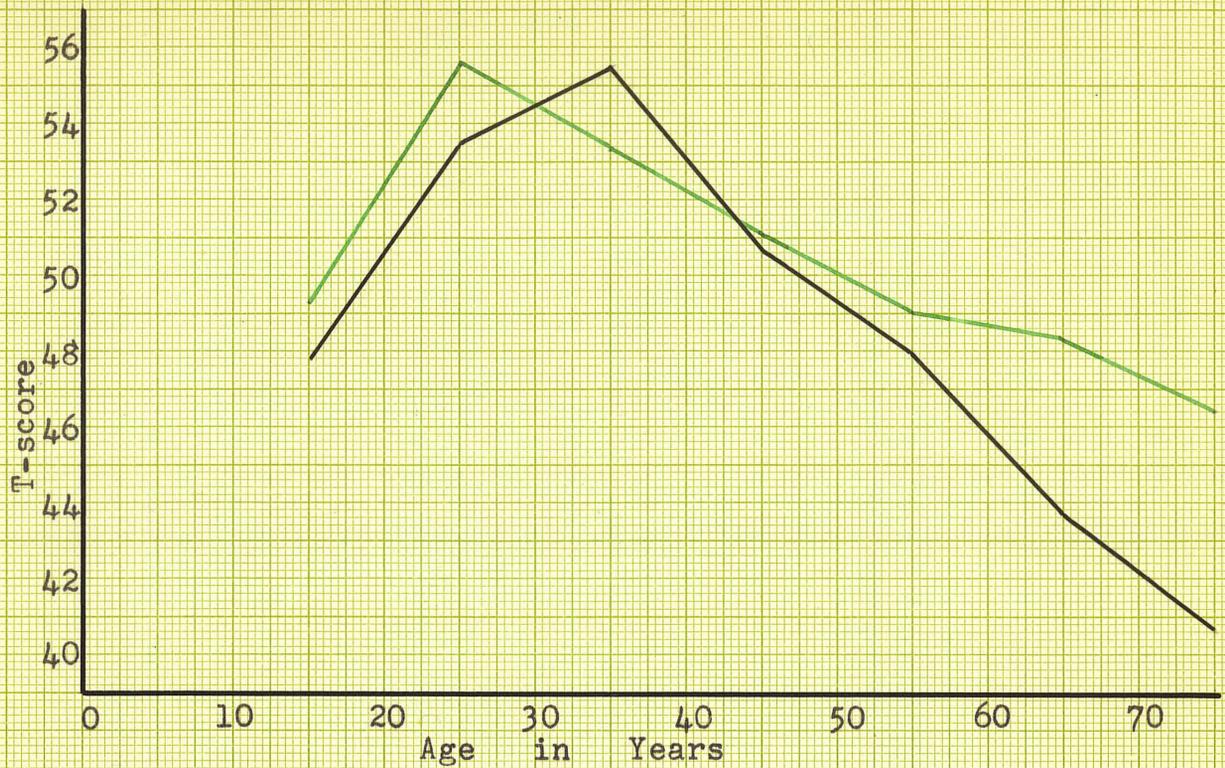


Figure 3. — Age changes in deductive thinking.  
— Age changes in ability to interpret data and weigh evidence.

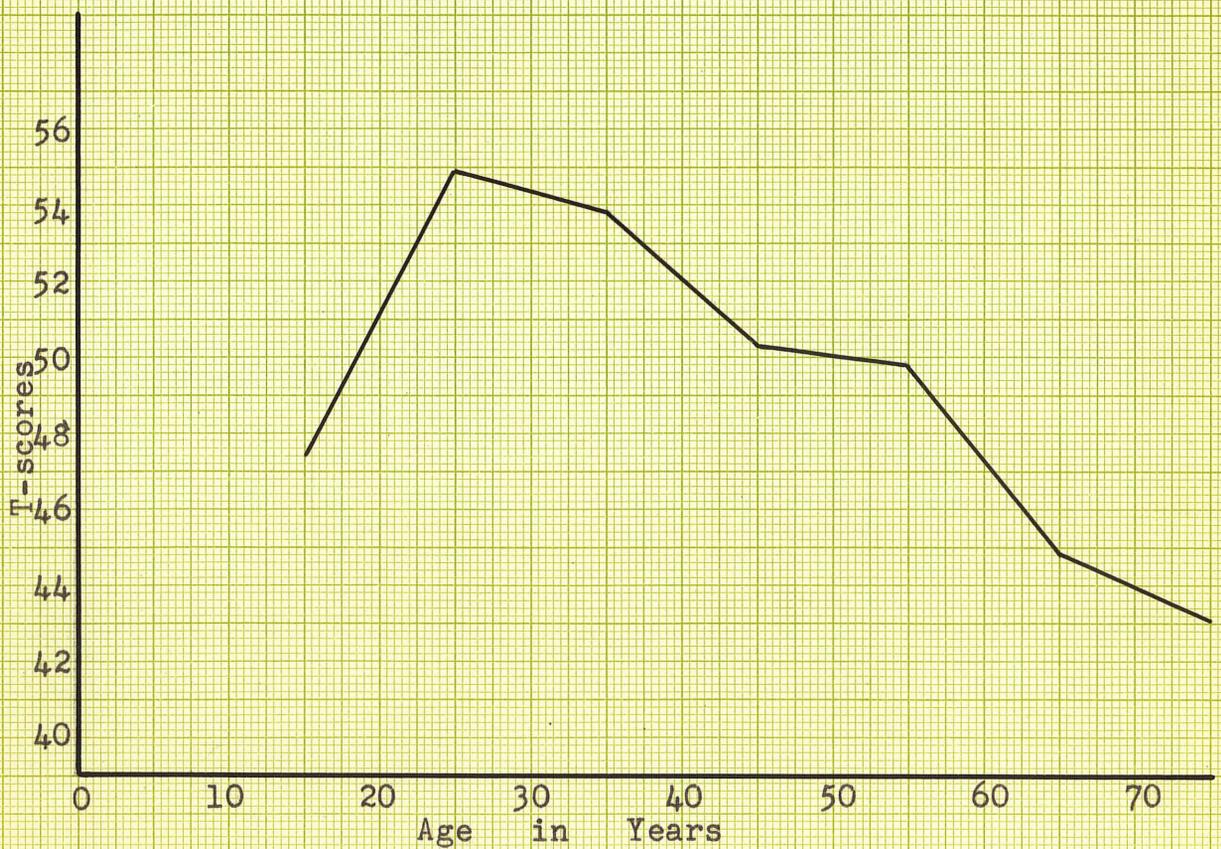


Figure 4. Age changes in ability to evaluate arguments.

Ability to draw inferences. Figure 2 indicates that the peak of ability to draw inferences appears to be at about age 35, after which it declines sharply. Between the ages of 45 and 65, however, there is a steadier maintaining of scores than there is by the over-65 group, which appears to lose this ability extremely rapidly.

Ability to detect assumptions. The peak mean for the ability to recognize assumptions has been closely approached by the age of 25 and is maintained till 35, as shown in Figure 2. The decline of this ability appears to be less sharp thereafter, than the decline of inferential abilities. Nevertheless, it is steady throughout each of the successive age groups.

Deductive thinking. Figure 3 shows that the peak of ability to make deductions is arrived at by age 35, and that it is approached much more gradually than the preceding sub-ability peaks for drawing inferences and recognizing assumptions. After 35 the decline is steady till the age of 55, after which the rate of decline is fairly rapid.

Ability to interpret data. It appears from Figure 3 that ability to interpret data and weigh evidence achieves its peak earlier than that of the preceding three subabilities, being reached at age 25. There is no stage of maintenance

for this subability, but a decline at an even rate towards the later years.

Ability to evaluate arguments. Figure 4 indicates that, as with the ability to interpret data, the ability to evaluate arguments approaches its peak at 25 years. The rate of decline is approximately the same. Although not quite so steady, it holds up fairly well between 45 and 55.

#### Individual Differences

The preceding results portray a very gloomy picture as far as later age changes in critical thinking are concerned. It is important to remember, however, that these conclusions are based on the mean performance of subjects in different decades, with no attention being given to individual variations in test performance. The picture becomes much more encouraging when one determines the percentage of subjects in the later decades who obtain scores equal to or greater than those obtained by subjects at their peak, that is, those in their twenties and thirties. Table IV shows such an analysis. In it is given the percentage of subjects in the different decades receiving scores equal to or greater than the peak mean of 70.34 (raw score on total test) obtained by the 20 to 39 age group. It can be seen that, while the percentage of subjects achieving scores over the peak mean declines with age, there is still a fairly high

percentage of individuals in the older age group who obtain scores as high or higher than this peak. For example, we find that approximately 16 per cent of subjects in the sixties do as well or better than those in their youth; about 21 per cent of those in the fifties do as well or better, and approximately 28 per cent of those in the forties do as well or better than those at their peak. Thus, when individual differences are taken into consideration, the decrement in critical thinking in the later years is not as pronounced as it appears to be at first sight.

TABLE IV

PERCENTAGE OF SUBJECTS IN DIFFERENT DECADES OBTAINING SCORES EQUAL TO OR ABOVE THE PEAK MEAN OF 70.34 OBTAINED BY THE 20 - 39 AGE GROUP

Age Group	Number of Subjects	Raw Score Mean	Number over Mean of 70.34	Percentage over Mean of 70.34
40's	82	65.09	23	28.05
50's	67	62.19	14	20.89
60's	57	58.05	9	15.79
Over 70	35	53.31	3	8.57

#### Rigidity vs. Flexibility in Critical Thinking

During the course of administration of the tests

it was observed that certain subjects, usually the older ones, often exhibited a high degree of mental rigidity in making critical judgments. In order to ascertain whether any significant difference in rigidity or flexibility in thinking existed between the various age groups, the scores for Test 1, Inference, were examined. In this subtest, the subject is required to make a choice between different degrees of truth and falsity. The items can be judged true, probably true, false, probably false, or containing insufficient data for making a correct decision. For purposes of statistical analysis, the 484 subjects were divided into three age groups: under 25, 35 to 45, and over 60 years.

Table V shows the mean number of choices made by the three age groups of various possible answers to questions requiring inferences. It can be seen from this table that the mean score for the extreme choices of true or false increases with age, from 2.97 to 4.15. This increase is statistically significant ( $t = 8.43$ ,  $p < .001$ .) Table V also shows that the number of "insufficient data" responses decreases with age (5.60 vs. 4.23) as does also the mean number of "probably true or false" answers (4.17 vs. 3.63.) Although the first difference is statistically significant ( $t = 2.69$ ,  $p = .01$ ), the second is not ( $t = 1.72$ ,  $p > .05 < .10$ .)

From these results it would appear that the older subjects are more inclined to choose the absolutes of true and false than to consider the intermediate degrees of truth and falsity, or to consider the possibility of insufficient data being provided. This relative inflexibility on the part of the older subjects is undoubtedly one of the reasons for the decrement in critical thinking ability which occurs after the mid-thirties.

TABLE V

MEAN NUMBER OF CHOICES BY THREE AGE GROUPS, OF FIVE POSSIBLE ANSWERS TO QUESTIONS REQUIRING INFERENCE

Age Group	Number of Subjects	Mean Number of Choices		
		True and False	Probably True or False	Insufficient Data
Under 25	110	2.97	4.17	5.60
35 - 45	94	3.61	3.33	5.90
Over 60	92	4.15	3.63	4.23

#### Objectivity in Critical Thinking

In a further attempt to discover the reasons for the fairly pronounced decrement in critical thinking with age, an analysis was made of the objectivity with which subjects of different ages answered the problems posed in the test.

An index of objectivity could easily be devised, since the test problems are such that they require critical thinking about two different kinds of subject matter. In half the items, the subject is asked to think critically about problems involving "neutral" subject matter such as the weather, scientific facts or experiments, and other things concerning which people generally do not seem to have strong emotional feelings or prejudices. The other half of the test is composed of items that are approximately parallel in logical structure, but pertain to subject matter involving political, economic, social and racial issues toward which people are apt to have emotional feelings, biases or prejudices. The inclusion of subject matter from various areas of controversy or prejudice should provide a partial sample of a person's thinking about issues regarding which he is apt to have personal biases. Thus, any subject's total critical thinking score is likely to be reduced by any lack of objectivity in his thinking about the problem posed.

An index of objectivity for each subject was calculated by dividing his score on the "emotionally loaded" items by his score on the "neutral" items, and multiplying by 100. Before doing this, however, it was necessary to identify those items which would most clearly have an emotional loading for the particular group involved, that is, the local Canadian sample, and those items which would have

no such emotional loading for the local group studied. This factor is particularly important since the Watson-Glaser test of critical thinking is an American test employing questions which are emotional or neutral for Americans but not necessarily so for Canadians. This identification of the two types of items was achieved by presenting the 99 items of the test to a group of thirteen carefully picked local judges. These judges were required to rate each item on a three-point scale with the following criteria:

2. Strong emotional bias. This rating applies if many people are likely to feel strongly one way or another about the item. Without a doubt, the item would be decided on the basis of previous experiences of a strong emotional nature with similar material or situations.
1. Mild emotional bias. These items are not considered strongly loaded with emotional prejudice or bias, but could give rise to occasional or moderate emotional defensive reactions or feelings. They may have a stronger emotional loading for certain individuals, but not for majority groups or for the population generally.
0. Neutral items, which are relatively devoid of emotional connotation, except for occasional atypical individuals.

The mean weights for the 99 items were then computed. Those items receiving the highest weights were considered to be "emotionally toned" while those showing the lowest weights were considered to be the most "neutral." A total

of 30 items (15 of each) was thus selected. These items are listed in Table VII, Appendix "E."

For purposes of statistical analysis the 484 subjects were divided into three age groups: under 25 years, 35 to 45 years, and over 60 years. For each subject in these age groups an objectivity index was calculated in the manner mentioned previously (emotional/neutral.) In such an index, a high score indicates a high degree of objectivity in answering the test questions, and a low score, low objectivity or subjectivity. Table VI summarizes the results. It can be seen that as age increases

TABLE VI  
OBJECTIVITY INDICES FOR THREE AGE GROUPS,  
UNDER 25, 35-45 YEARS AND OVER 60 YEARS

Age Group	Number of Subjects	Objectivity Index
Under 25	110	102.445
35 - 45	94	99.521
Over 60	92	96.163

there is a progressive decrease in the objectivity with which the subjects answered the test items. The difference in scores between the under 25 age group and the over 60 group is statistically significant ( $t = 2.01$ ,  $p > .02 < .05$ .)

From this it would appear that the over 60 group had allowed their emotional feelings, biases and prejudices to reduce the objectivity of their thinking. This undoubtedly is another factor partly responsible for the large decrement in critical thinking shown by the elderly subjects.

## II. DISCUSSION OF RESULTS

The preceding results have indicated that critical thinking ability is poorly developed in childhood and that it does not reach the peak of development until the mid-twenties. Following this, it holds steady for some years, showing a gradual and progressive decline after the age of 35, through to the period of old age. This curve of development and decline is quite similar to those which have been obtained in studies of general intelligence by the use of such tests as the Army Alpha (42), the Otis (49), and the Wechsler-Bellevue (76). However, there is one main difference. Critical thinking ability matures later and begins to decline later than general intelligence which reaches its peak in the late teens or early twenties and then immediately begins to decline. This difference might be expected since critical thinking is an ability which presumably is dependent upon a considerable amount of accumulated experience.

The curve of development and decline of critical thinking ability, as demonstrated in this study, is also

quite similar to those obtained in three other studies of intellectual processes. For example, Raven (see Figure 80 of Zubek and Solberg, 80, p. 270), studying subjects ranging in age from 6 to 60 years, found that the "capacity to understand and apply a fresh method of thinking" reaches a peak in the mid-twenties and then slowly declines thereafter.

A similar curve of development and decline of the higher thought processes was the result of Bromley's study in 1956 (8). He administered the Shaw test to 256 "relatively normal" adults divided into four age groups (17-35, 35-51, 51-66, 66-82.) This test is one of creative intellectual output and involves the mental operations of abstracting, serializing and productive thinking, and personal qualities of persistence, flexibility, originality and fluency. Bromley found that the peak years for both quality and quantity of output occurred in the youngest group (mean age, 27 years) and that both quality and quantity declined with age, quality of output declining more rapidly.

Lehman's study (47) of age and achievement resulted in a growth curve which showed a rapid rise from the age of 20 to the early thirties, followed by a slow decline to the 45-50 interval, with maximum productivity at 33 and a rapid drop after the age of 55. Lehman concluded, however, that while man's most creative years occur relatively early in

life, and quality of achievement reaches its peak earlier and declines more rapidly than mere quantity achievement, very high qualitative achievement reaches its peak earlier and declines more rapidly than achievement of lesser merit. The peak year for both quality and quantity of achievement in this study was shown to be 27.

All of these studies seem to indicate that the overall curve of development and decline is essentially the same for the various kinds of intellectual processes.

In the present investigation, the pattern of development and decline of the five subabilities of critical thinking was shown to be anything but uniform. The ability to interpret data and the ability to evaluate arguments both reach a peak in the mid-twenties and decline almost immediately. On the other hand, the subabilities of drawing inferences, detecting assumptions and making deductions do not reach their peaks till some ten years later, after which they also begin to decline. Furthermore, the rate of decline of the five subabilities is shown to vary considerably. Other investigators have also found a similar differential pattern of growth and decline in subabilities. Jones and Conrad (42), for instance, found such differences occurring on the individual subtests of the Army Alpha. Scores on verbal analogies, they pointed out, declined more quickly than those on arithmetical reasoning, and the latter

more quickly than those on general information. This study also revealed a fairly marked decline in the old age groups, of their ability to form comparisons, to reason by analogy, exercise judgment, plan for the future, detect similarities and to do arithmetic problems. Eysenck (21) found that the most rapid decline was on tests deducing relations and correlates. It would seem clear, then, that the specific effect of age on intellectual processes is a function of the particular abilities being studied.

This discussion has so far dealt with the overall or mean performance of subjects in the various decades and the results indicate that older subjects perform much worse than younger ones. The use of mean scores, however, can be quite deceptive, since it fails to show the great overlap in score distribution from one decade to the next. This overlap is frequently evident in such investigations as the present one. Here, it was shown that approximately sixteen per cent of the subjects in their sixties and twenty-one per cent in their fifties do as well or even better than those at the peak of their ability (see Table IV.) Essentially the same results were obtained by Miles (50, 51) in the Stanford University Studies in Later Maturity, who reported that anywhere from 10 to 25 per cent of those in later

maturity and old age do as well or better than those in middle maturity or at their peak.

The question arises, then, as to why a decrement in performance occurs in later years. Several reasons are offered when such results are revealed in experiments on aging. One of these is the speed, or time factor--the older person cannot work as quickly as a younger subject and consequently receives a lower score. This factor was not operative in the present experiment, since a power test was employed.

A second reason for the occurrence of a decrement in old age performance is that of poor or faulty motivation on the part of the older subjects. It is doubtful whether this is a significant factor in the present investigation. The subjects were all volunteers and no pressure of any kind was used to secure cooperation. Any individual showing reluctance or indifference was rejected. On the other hand, the subjects showed a keen interest in solving the problems. This enthusiasm on the part of many of the subjects may be due largely to the nature of the test problems. These were mostly of a realistic or practical nature, dealing with political, economic, social and scientific issues, all of which elicit the interest of most people.

A third reason offered for explanation of decrement is that the decline is attributable to a lack of recent

practice in the abilities tested. Again, this factor is not believed to have played too great a role in accounting for the present results. The problems were of the kind that people often encounter in daily life. Furthermore, since most of the older subjects were in the professional or semi-professional categories, they of necessity had to use their critical abilities constantly, in some degree, in the course of their normal duties.

An item analysis of the subtests of the Critical Thinking Appraisal Test revealed that some of the decrement is most likely due to the operation of two factors. The first is low objectivity in dealing with certain types of problems. It was found that the older subjects had allowed their emotional feelings and prejudices to reduce the objectivity of their thinking when dealing with racial, social, political and other issues about which people have strong personal feelings.

The second possible factor in decline is that of mental "rigidity." Older subjects, in answering certain of the test problems, tended to choose the absolutes of true and false, rather than to consider other possible alternative answers. This "rigidity" on the part of the older subjects in choosing among alternatives is a well-known phenomenon in the literature on gerontology. For example, Heglin (36) used the Luchins Water Jars test to investigate the operation

of "set" in different age groups. This test measures a subject's degree of rigidity or flexibility in solving problems. He reported that the older group (50 to 85 years) showed more set on all measures, e.g., susceptibility to set, ability to surmount set, and trainability in avoiding set. The middle group (20 to 49 years) showed less, and the younger group (14 to 19 years) the least set. The older group improved least with training. Similar results were reported by Anderson and Dvorak (2) who asked a group of subjects to evaluate various moral issues according to a number of standards such as absolute standards of right and wrong, standards of public opinion and standards set by intelligent judgment. They reported that a very high percentage of middle-aged people made their judgments according to absolute criteria. Similarly, Jones (43), using teachers of different ages, found that the older ones tended to classify various issues according to absolute standards.

In addition to these two factors of low objectivity and mental rigidity, physical changes in the central nervous system undoubtedly also play a part in producing the decrement in critical thinking in old age. Zubek and Solberg (80) have pointed out a number of these degenerative changes. There is, for instance, a consistent pattern of decrement in the size and weight of the brain after 60 years. There are

also changes in pigmentation, such as the loss of staining properties in the cell nucleus and the infiltration of pigments and fatty substances into the nerve cells. Other investigators have reported, with considerable consistency, the complete disappearance (by neuronophagia) of some cells in the brain. There are also reports of the greater incidence of slow brain waves after the age of 65, throughout all parts of the cerebral cortex. These EEG changes indicate a depression of neural activity and are undoubtedly related to the decline of psychological processes.

From such findings as these, then, it would seem that both psychological factors (such as emotional and attitudinal) and physical factors are responsible for the decrement in critical thinking abilities which appears in the later years of life.

## CHAPTER IV

### SUMMARY AND CONCLUSIONS

This investigation has dealt with the development and decline of one particular type of higher intellectual process, viz. critical thinking ability. Previous research has been extremely meagre and has dealt for the most part with limited areas of development. Studies have been confined, for example, to that of establishing the age of the beginnings of thinking processes in young children, or of comparing the performance of older and younger subjects in tests of conceptual thinking, deductive and inductive reasoning, the use of logic and creative thinking, all of which are involved in some degree in critical thinking. But none of these can be said to be, in itself, a measure of critical thinking ability. Furthermore, very few of the studies have been concerned with the life span changes in these abilities.

Recently, several investigators have subjected this particular ability to examination and have constructed a reasonably good measure of it. This is the Watson-Glaser Critical Thinking Appraisal test. In the present investigation, this test was administered to 484 subjects ranging in age from 12 to 89 years. Within each age group there was an equal representation of the sexes and a diversification of

educational, occupational and economic backgrounds.

Raw scores obtained for each of the five subtests, inference, recognition of assumptions, deduction, interpretation and evaluation of arguments, were converted to T-scores and graphs representing the developmental growth and decline of each of these five subabilities of critical thinking were drawn up.

From the data obtained from each of five subtests of critical thinking ability and from a composite of these five subabilities, the main findings are as follows:

Critical thinking ability shows a progressive increase from late childhood through to the mid-twenties, holds up to 35 years and then declines progressively through to the seventies. The five subabilities of critical thinking show a differential pattern of development and decline. Although the older subjects as a whole scored poorly in critical thinking, many of them still showed a high level of performance--a performance as high or higher than that of young adult subjects. The poor performance in critical thinking ability in the older subjects is partly due to the two factors of rigidity in thinking, and low objectivity. In answering the test problems, the older subjects as a whole were inclined to choose the absolutes of true or false, rather than to consider intermediate degrees of truth and falsity or other alternatives. The older subjects

also allowed their emotional feelings and biases to reduce the objectivity of their thinking about certain problems.

A P P E N D I C E S

A P P E N D I X "A"

# WATSON-GLASER

## CRITICAL THINKING

### APPRAISAL

*Form BM*

---

by GOODWIN WATSON

*Professor of Education, Teachers College,  
Columbia University*

and EDWARD MAYNARD GLASER

*Consulting Psychologist, Rohrer, Hibler, and Replogle,  
Los Angeles*

#### DIRECTIONS:

This booklet contains several different types of tests designed to find out how well you are able to reason analytically and logically.

**Do not turn this page until instructed to do so.** Do not make any marks on this test booklet. All answers are to be marked on the separate Answer Sheet provided. If you wish to change an answer, be sure to erase your old answer completely.

### TEST 1. Inference

**DIRECTIONS.** An *inference* is a conclusion which a person draws from certain observed or supposed facts. Thus, from the electric light visible behind the window shades and from the sound of piano music in a house, a person might *infer* that someone is at home. But this inference may or may not be correct. Possibly the people in the house went out leaving the lights on, and the piano music could be coming from a radio or phonograph they left playing.

In this test each exercise begins with a *statement of facts* which you are to regard as true. After each statement of facts you will find several possible inferences — that is, inferences which some persons might make from the stated facts. Examine each inference separately, and make a decision as to its degree of truth or falsity.

On the Answer Sheet you will find for each inference spaces marked with the letters T, PT, ID, PF, and F. For each inference make a mark on the Answer Sheet under the appropriate letter as follows:

- T** — if you think the inference is *definitely true*; that it properly follows from the statement of facts given.
- PT** — if, in the light of the facts given, you think the inference is *probably true*; that there is better than an even chance that it is true.
- ID** — if you decide that there are *insufficient data*; that you cannot tell from the facts given whether the inference is likely to be true or false.
- PF** — if, in the light of the facts given, you think the inference is *probably false*; that there is better than an even chance that it is false.
- F** — if you think the inference is *definitely false*; that it cannot possibly be drawn from the facts given or that in some manner it contradicts the facts.

Sometimes, in deciding whether an inference is probably true or probably false, you will have to use certain commonly accepted knowledge or information which practically every person knows. This will be illustrated in the example which follows.

Here is the example; the correct answers are indicated in the block at the right.

**EXAMPLE.** A thousand eighth-grade students recently attended a voluntary week-end conference in a Midwestern city. At this conference questions of race relations and means of achieving lasting world peace were discussed, since these were the problems the students felt to be most vital today.

1. As a group, the students who attended this conference had a keener interest in humanitarian or broad social problems than most eighth-grade students have.....
2. The majority of these students were between the ages of 17 and 18.....
3. The students came from all sections of the country.....
4. The students came to discuss trade-union problems.....
5. Some eighth-grade students felt that discussion of race relations and means of achieving world peace might be worthwhile.....

#### TEST 1

	T	PT	ID	PF	F
1	•••	█	•••	•••	•••
2	•••	•••	•••	█	•••
3	•••	•••	█	•••	•••
4	•••	•••	•••	•••	█
5	█	•••	•••	•••	•••

In the above example, inference 1 is *probably true* (PT) because (as is common knowledge) most eighth-grade students are not likely to evidence such serious concern with broad social problems.

Inference 2 is *probably false* (PF) because (common knowledge) there are relatively few eighth-grade students in the United States between 17 and 18 years of age.

There is no evidence for inference 3. Thus there are *insufficient data* (ID) for making a judgment in this matter.

Inference 4 is *definitely false* (F) because it is given in the statement of facts that race relations and means for achieving world peace were the problems discussed.

Inference 5 necessarily follows from the given facts; it therefore is *true* (T).

In the exercises which follow, more than one of the inferences from a given statement of facts may be true (T) or false (F), or probably true (PT), or probably false (PF) or have insufficient data (ID) to warrant any conclusion. That is, you are to consider each inference by itself.

Make a heavy black mark in the space under the letter that you think best describes each inference. If you change an answer, erase thoroughly. Make no extra marks on the answer sheet.

A study in the Elmtown High School showed that the students from high-income homes received better grades on the average than did pupils from low-income homes. Differences in school grades given by teachers were considerably greater than differences on standard tests of intelligence and achievement. Students from high-income homes took part in many more of the extracurricular school activities which cost money than did the students from low-income homes. Student officers were usually chosen from the high-income group.

1. Elmtown sent children of rich and poor to the same high school.....
2. Many students from low-income homes felt they couldn't afford to participate in extracurricular school activities which cost money.....
3. The students from low-income homes actually studied harder than did the students from wealthy homes. ....
4. There were no differences between students from high-income homes and students from low-income homes on standard tests of intelligence and achievement.....
5. The majority of the students from low-income homes, recognizing the general superiority of the students from high-income homes, wanted them to lead the student groups.....

---

The town of Westfield, beginning twenty years ago, as gradually bought up farms abandoned by owners who failed to pay taxes, and by this time has set out some 3600 acres of community forest. The oak trees have grown rapidly. The town forests yielded \$9000 net profit on lumber last year and \$8500 the year before. Local authorities believe that the net profit on the lumber will eventually be \$40,000 a year.

6. It costs the town more to cut and sell the lumber than it realizes from the sales.....
7. The operating expenses were less than \$9000 last year.....
8. The owners who abandoned their farms and failed to pay their taxes were either incompetent farmers or lazy ones.....
9. The town of Westfield is continuing to buy up tax-delinquent farms to be set aside as community forests.....
0. The Westfield community forests will yield an annual net profit of \$40,000 within two or three years.....

When a Negro doctor bought a home in a suburban residential district where only whites of various nationalities had lived before, a mob stoned the house and broke many windows. The mayor, at the request of his Committee on Race Relations, sent extra police to keep order in this district.

11. The immediate neighbors of the Negro doctor liked him and did not join the mob.....
12. The Negro doctor was a better citizen than the white people who lived around him.....
13. The mob was composed of citizens who upheld democratic principles.....
14. The Negro family sold the house and moved elsewhere.....
15. The white families in the district all quickly sold their homes and moved away.....

---

Mr. Brown, who lives near the town of Salem, was brought before the Salem municipal court for the fourth time in the past month on a charge of keeping his dance hall open after midnight. He again pleaded guilty and was fined the maximum, \$100, as in each earlier instance.

16. Mr. Brown thought that it would pay him financially to keep his place open after 12 o'clock, even though he had to risk paying frequent fines.....
17. Mr. Brown's dance hall is within the legal jurisdiction of the town of Salem.....
18. The regular patrons of Mr. Brown's dance hall are habitual lawbreakers.....
19. The maximum fine of \$100 was fully effective in keeping all dance halls in Salem closed after midnight.
20. The midnight closing law was enacted because one of Mr. Brown's neighbors complained about the noise.....

---

*Go on to the next page.*

TEST 2. *Recognition of Assumptions*

DIRECTIONS. An *assumption* is something supposed or taken for granted. When someone states, "I'll graduate in June," he takes for granted or assumes that he will be alive in June, that he will remain in school until that time, that he will pass his courses, and similar things.

Below are a number of statements. Each statement is followed by several proposed assumptions. *You are to decide for each assumption whether it necessarily is taken for granted in the statement.*

If you think the given assumption is taken for granted in the statement, make a heavy mark between the dotted

lines under "ASSUMPTION MADE" in the proper place on the Answer Sheet. If you think the assumption is not necessarily taken for granted in the statement, make a heavy line under "ASSUMPTION NOT MADE" on the Answer Sheet.

Below is an example: the block at the right shows how these items should be marked on the Answer Sheet.

If you do not see why the answers marked are right, ask the examiner to explain. In some cases more than one of the given assumptions is necessarily made; in other cases none of the given assumptions is made.

**EXAMPLE.** STATEMENTS: "We need to save time in getting there, so we'd better go by plane."

PROPOSED ASSUMPTIONS:

1. Going by plane will take less time than going by some other means of transportation. (It is assumed in the statement that greater speed of a plane over other means of transportation will enable the group to get to their destination in less time.)
2. It is possible to make plane connections to our destination. (This is necessarily assumed in the statement, since, in order to save time by plane, it must be possible to go by plane.)
3. Travel by plane is more convenient than travel by train. (This assumption is not made in the statement — the statement has to do with saving time, and says nothing about convenience or about any other specific mode of travel.)

TEST 2	
ASSUMPTION MADE	NOT MADE
1 ■	⋮
2 ■	⋮
3 ⋮	■

STATEMENT: "The coach isn't going to let Ed play in any more games this season."

PROPOSED ASSUMPTIONS:

21. The coach can prevent Ed from playing.
22. Ed must have done something terribly wrong.
23. The games for this season are not yet over.
24. There will be a chance for Ed to play next season.

STATEMENT: "The less government interferes with business, the better for everyone."

PROPOSED ASSUMPTIONS:

25. Government is inefficient.
26. Most businessmen are superior in character to most government officials.
27. Any good cause is injured if it gets mixed up in politics.
28. If free from government controls, business would do what is good for us.

STATEMENT: "I want to be sure I don't get typhoid fever while I'm in South America, so I shall go to my physician and get typhoid injections before I sail."

PROPOSED ASSUMPTIONS:

29. If I don't take the injections, I shall become ill with the fever.
30. Only a physician can give typhoid injections.
31. I am going to South America very soon.

STATEMENT: "Only a fool would try to keep the pleasant taste of wine by having his mouth always full of it."

PROPOSED ASSUMPTIONS:

32. Some fools don't like wine.
33. A wise man never has his mouth full of wine.
34. Some fools get drunk on wine.

STATEMENT: "If war is inevitable, we'd better launch a preventive war now while we have the advantage."

PROPOSED ASSUMPTIONS:

35. If we fight now, we are more likely to win than we would be if forced to fight later.
36. War is inevitable.

Go on to the next page

TEST 3. *Deduction*

DIRECTIONS. Each exercise below consists of two statements (premises) followed by several proposed conclusions. For the purposes of this test, consider the two statements in each exercise as true without exception. Read the first conclusion beneath the statements, and if you think it *necessarily follows from the statements given*, answer by making a heavy black mark between the pair of dotted lines under "CONCLUSION FOLLOWS" in the corresponding blank on the Answer Sheet. If you think it is *not a necessary conclusion* from the given statements,

then put a heavy black mark under "CONCLUSION DOES NOT FOLLOW," even though you may believe it to be true from your general knowledge.

Likewise read and judge each of the other conclusions. Try not to let your prejudices influence your judgment — just stick to the given statements and judge each conclusion as to whether it necessarily follows from them. Mark all your answers on the Answer Sheet.

Here is an example; the block at the right shows how your answers should be marked on the Answer Sheet.

**EXAMPLE.** Some holidays are rainy. All rainy days are boring. Therefore —

1. No clear days are boring. (The conclusion does not follow, as you cannot tell from these statements whether or not clear days are boring and some may be.) .....
2. Some holidays are boring. (The conclusion necessarily follows from the statements, since, according to them, the rainy holidays must be boring.) .....
3. Some holidays are not boring. (The conclusion does not follow from the statements even though you may know that some holidays are very pleasant.) .....

TEST 3	
CONCLUSION FOLLOWS	CONCLUSION DOES NOT FOLLOW
1 : : :	█
2 █	: : :
3 : : :	█

All good typists must be able to spell correctly. Anne can spell correctly. Therefore —

7. Anne is not a good typist.....
8. Anne's typing ability cannot be inferred from the information given.
9. Anne is a good typist.....

All glass is transparent. Lucite is not glass. Therefore —

0. Lucite is transparent.....
1. Some forms of glass are not transparent.....
2. Lucite is not transparent.....

All radicals are foreign-born. No patriotic citizen is a radical. Therefore —

3. No radical is a patriotic citizen.....
4. Liberals and Democrats work together for the same ends.....
5. No patriotic citizen is foreign-born.....
6. Some foreign-born people are patriotic citizens.....

Some Russians would like to control the world. All Russians seek a better life for themselves. Therefore —

7. Some people who would like to control the world seek a better life for themselves.....
8. All people who seek a better life for themselves would like to control the world.....
9. Many Russians find life under their dictatorship miserable, and therefore seek a change.....

Some students keep late hours. All students hate to get up in the morning. Therefore —

50. All students who hate to get up in the morning were up late the night before.....
51. No student likes to get up in the morning.....
52. All students who keep late hours hate to get up in the morning.....
53. Keeping late hours causes students to hate to get up in the morning.....

If Negroes are segregated, there is racial discrimination. If democratic principles are not violated, Negroes are not segregated. Therefore —

54. Racial discrimination is limited to Negroes.....
55. If there is racial discrimination, Negroes are segregated.....
56. If democratic principles are violated, there is racial discrimination.....

All children who are spanked develop repressions. John never was spanked. Therefore —

57. John has never developed repressions.....
58. All people who have repressions probably were spanked while children.....

If a person thinks straight, he will define the problem to be solved. Some people do not define the problem to be solved. Therefore —

59. If a person does not think straight, he will not define the problem to be solved.
60. If a person defines the problem to be solved, he thinks straight.
61. Some people do not know how to define the problem to be solved.

TEST 4. Interpretation

DIRECTIONS. Each exercise below consists of a short paragraph followed by several proposed conclusions.

For the purpose of this test assume that everything in the short paragraph is true. The problem is to judge whether or not each of the proposed conclusions logically follows beyond a reasonable doubt from the information given in the paragraph.

If you think that the proposed conclusion follows beyond a reasonable doubt (even though it may not follow

absolutely and necessarily), then make a heavy black mark between the appropriate dotted lines under the "CONCLUSION FOLLOWS" column on the Answer Sheet. If you think that the conclusion does not follow beyond a reasonable doubt from the facts given, then make a mark under "CONCLUSION DOES NOT FOLLOW."

In some cases more than one of the proposed conclusions may follow; in other cases none of the conclusions may follow.

A salesman for Brown's Liniment claimed that his product would promptly soothe sore muscles in the body because it would penetrate very quickly to the affected parts. In order to demonstrate the penetrating qualities of Brown's Liniment, the salesman poured ten drops on a thick piece of sole leather, and the liniment quickly went through this substance.

- 62. The liniment which the salesman poured on the sole leather was effective in penetrating the kind of sole leather used in the demonstration.
63. The salesman deliberately misrepresented his product.
64. The salesman's demonstration was good evidence for his claim that the liniment would promptly soothe sore muscles in the body.
65. There is no relationship between the liniment's ability to penetrate this particular piece of sole leather and its ability to penetrate into the human body.

In a certain city where school attendance laws are rigidly enforced, it was found that only 15 per cent of the male school population had a perfect attendance record during any single school semester. Among those who sold newspapers, however, 25 per cent had a perfect attendance record during the same period.

- 66. If truants were given jobs selling newspapers, their school attendance would improve.
67. Newsboys in another city would have a similarly superior attendance record.
68. Those who carry partial responsibility for their own support tend to take their schooling more seriously.

Among the people listed in Who's Who in America those with college degrees are about 15 times more numerous than those with only a grade-school education.

- 69. If the given figures hold constant, college graduates as a group have a 15 to 1 better chance of being listed in Who's Who than have people with only a grade-school education.
70. The editors of Who's Who give undue weight to the types of achievement made by college-trained people.
71. People listed in Who's Who tend to come from wealthy families who can afford to give their children many advantages, including a college education.

Jane's posture used to be poor; she dressed in bad taste, had very few friends, was ill at ease in company and in general was quite unhappy and maladjusted. Then someone recommended that she visit Dr. Baldwin a reputed expert on helping people to improve their personality. Jane took this recommendation, and after three months of treatment by Dr. Baldwin she carried herself well, dressed attractively, was more at ease and more popular, and in general felt much happier.

- 72. Jane's improvement may be only temporary.
73. Jane's marked improvement must have been brought about by important factors other than Dr. Baldwin's treatment.
74. Jane's improvement was caused solely by the treatment given her by Dr. Baldwin.

Several studies have shown that Southern Negroes make lower scores on intelligence tests than Northern Negroes, but that the average measured intelligence of Southern Negro children who move North at an early age increases each year until their average measured intelligence approaches that of the Northern Negroes.

- 75. The children of Southern Negroes who move North tend to equal the children of Northern Negroes in the kinds of ability measured by the intelligence tests used.
76. The increase is due to the relative superiority of the schools Negroes attend in the North.

TEST 4. Interpretation (Continued)

A national weekly magazine published some articles criticizing the action of the Catholic Church in matters of health and censorship, and was promptly banned from the high school libraries by the school board of an Eastern city.

- 77. The majority of the people on that school board were afraid of the power of the Catholic Church. . . .
- 78. A majority of the people in that city must have been Catholics. . . . .
- 79. The magazine should not have published those articles. . . . .

Statistics for a certain city in the United States, which has a well-organized Boy Scout program, indicate that during the past ten years no Scouts have been convicted of juvenile delinquency.

- 80. The Scout program prevents delinquency. . . . .
- 81. Scouting is a good way to meet the most important needs of boys today. . . . .

A sleeping wife was awakened by the sensation of a sharp blow across her mouth. Later that night her husband returned from a fishing trip, his front teeth knocked out by a sharp blow from the oar of his boat. Apparently the husband's accident and the wife's dream came at about the same time.

- 82. Under the circumstances described there was no way for any ordinary form of direct communication between the husband and wife at the time of the husband's accident. . . . .
- 83. If the facts reported in this case are true, they can be accounted for only by the operation of mental telepathy. . . . .
- 84. The dream was a chance coincidence which was not really influenced by the accident. . . . .
- 85. There are many occurrences in life which we do not understand. . . . .

Go on to the next test.

TEST 5. Evaluation of Arguments

DIRECTIONS. In making decisions about important questions it is desirable to be able to distinguish between arguments that are strong and those which are weak insofar as the question at issue is concerned.

Strong arguments must be both important and directly related to the question.

Weak arguments may not be directly related to the question, even though they may be of great general importance; or they may be of minor importance; or they may be related to trivial aspects of the question.

Below is a series of questions. Each question is followed by three or four arguments. For the purpose of this test you are to regard each argument as true. The problem then is to decide whether it is a STRONG argument or a WEAK argument.

You are to answer by making a heavy mark on the answer sheet under "STRONG" if you think the argument is strong, or by making a heavy mark under "WEAK" on the Answer Sheet if you think the argument is weak. When evaluating an argument, judge it on its own merit; try not to let counter-arguments or your own attitude toward the question influence your judgment. Judge each argument separately. In some questions all the arguments may be STRONG, in others all may be WEAK.

Here is an example. The block at the right shows how these arguments should be marked on the Answer Sheet. Study them carefully until you know just what is expected of you. Note that the argument is evaluated as to how well it supports the side of the question indicated.

<p><b>EXAMPLE.</b> Should all young men go to college?</p> <p>1. Yes; college provides an opportunity for them to learn school songs and cheers. (This would be a silly reason for spending years of one's life in college.)</p> <p>2. No; most young men profit more from work experience than from college classes. (According to the directions, we must accept this argument as true; hence it is a strong and important one against all young men going to college.) . . . . .</p> <p>3. No; excessive studying permanently warps an individual's personality. (This argument, although of great general importance when accepted as true, is not directly related to the question, because attendance at college does not necessarily require excessive studying.) . . . . .</p>	<p><b>TEST 5</b></p> <p><b>ARGUMENT</b></p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"><b>STRONG</b></td> <td style="width: 50%;"><b>WEAK</b></td> </tr> <tr> <td>1 : : : <input checked="" type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td>2 <input checked="" type="checkbox"/> : : :</td> <td><input type="checkbox"/></td> </tr> <tr> <td>3 : : : <input checked="" type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </table>	<b>STRONG</b>	<b>WEAK</b>	1 : : : <input checked="" type="checkbox"/>	<input type="checkbox"/>	2 <input checked="" type="checkbox"/> : : :	<input type="checkbox"/>	3 : : : <input checked="" type="checkbox"/>	<input type="checkbox"/>
<b>STRONG</b>	<b>WEAK</b>								
1 : : : <input checked="" type="checkbox"/>	<input type="checkbox"/>								
2 <input checked="" type="checkbox"/> : : :	<input type="checkbox"/>								
3 : : : <input checked="" type="checkbox"/>	<input type="checkbox"/>								

Remember that for the purpose of this test each argument is to be regarded as true.

Should the government provide "baby bonuses" to help support each dependent child in a family so that the family standard of living is not lowered by having children?

- 86. Yes; every family then could afford proper child care, which would greatly improve the general health of the nation.....
- 87. No; such bonuses would require much higher taxes, thus actually lowering the general standard of living.....
- 88. No; people who do not wish to have children then would be taxed to help provide "baby bonuses" for other people.....

Should the government continue to pay farmers the cost of soil-conservation practices on their own land?

- 89. Yes; food is a necessity to the entire nation, and farmers are much more likely to undertake the necessary practices which will assure abundant future crops if they get paid for the extra work involved.....
- 90. No; soil conservation is simply good farming practice which will increase the owner's likelihood of making a good living from his land; to tax others to pay him for helping himself is needless and unjust.
- 91. No; to pay a special group such as farmers out of public funds for helping themselves financially is opposed to the general welfare of the people as a whole.....

Is it possible for man to develop a death ray that will kill all living beings on whom it is focused?

- 92. No; no one has ever seen such a death ray outside the comics.....
- 93. Yes; powerful concentrations of X rays and other types of radiant energy do kill animals and human beings.....
- 94. No; some very competent scientists doubt that such a death ray is possible.....
- 95. No; for if man ever does develop such a ray, he also will work on counter measures to offset it.....

Would a labor party be a good thing for the people of the United States?

- 96. Yes; differences between Republicans and Democrats today are not so great as differences between liberals and conservatives within those parties.....
- 97. No; labor unions have sometimes called strikes which hurt the public at large.....

Should pupils be excused from public schools to receive religious instruction during school hours in their own churches?

- 98. No; having children go back and forth between school and church is inefficient and would violate our basic and time-tested concept of the desirability of separation between church and state. ....
- 99. Yes; if children do not receive religious instruction during school hours, it may be difficult to get them to church at any other time.....

Go back and check your work

APPENDIX "B"

# WATSON-GLASER CRITICAL THINKING APPRAISAL

*Answer Sheet for Forms AM and BM*

Class \_\_\_\_\_ Sex \_\_\_\_\_ Age \_\_\_\_\_ Yrs. \_\_\_\_\_ Mos. \_\_\_\_\_ Days \_\_\_\_\_ City \_\_\_\_\_ County \_\_\_\_\_ State \_\_\_\_\_

BE SURE YOUR MARKS ARE HEAVY AND BLACK. ERASE COMPLETELY ANY ANSWER YOU WISH TO CHANGE.

TEST 1						TEST 2		TEST 3		TEST 4		TEST 5		
INFERENCE						ASSUMPTION		CONCLUSION		CONCLUSION		ARGUMENT		
	T	PT	ID	PF	F	MADE	NOT MADE	FOLLOWS	DOES NOT FOLLOW	FOLLOWS	DOES NOT FOLLOW	STRONG	WEAK	
1	:	:	:	:	:	21	:	:	:	37	:	:	86	:
2	:	:	:	:	:	22	:	:	:	38	:	:	87	:
3	:	:	:	:	:	23	:	:	:	39	:	:	88	:
4	:	:	:	:	:	24	:	:	:	40	:	:	89	:
5	:	:	:	:	:	25	:	:	:	41	:	:	90	:
6	:	:	:	:	:	26	:	:	:	42	:	:	91	:
7	:	:	:	:	:	27	:	:	:	43	:	:	92	:
8	:	:	:	:	:	28	:	:	:	44	:	:	93	:
9	:	:	:	:	:	29	:	:	:	45	:	:	94	:
10	:	:	:	:	:	30	:	:	:	46	:	:	95	:
11	:	:	:	:	:	31	:	:	:	47	:	:	96	:
12	:	:	:	:	:	32	:	:	:	48	:	:	97	:
13	:	:	:	:	:	33	:	:	:	49	:	:	98	:
14	:	:	:	:	:	34	:	:	:	50	:	:	99	:
15	:	:	:	:	:	35	:	:	:	51	:	:		:
16	:	:	:	:	:	36	:	:	:	52	:	:		:
17	:	:	:	:	:		:	:	:	53	:	:		:
18	:	:	:	:	:		:	:	:	54	:	:		:
19	:	:	:	:	:		:	:	:	55	:	:		:
20	:	:	:	:	:		:	:	:	56	:	:		:
										57	:	:		:
										58	:	:		:
										59	:	:		:
										60	:	:		:
										61	:	:		:

## APPENDIX "C"

RAW SCORES OBTAINED BY 484 SUBJECTS IN A  
TEST OF CRITICAL THINKING ABILITIESTeens (N = 63)

44	49	51	54	58	60	62	65	68	71	77
44	49	51	54	58	61	63	65	69	73	80
47	50	52	55	59	61	63	66	69	74	80
47	50	53	56	59	61	64	67	70	74	..
47	50	53	56	60	61	64	68	70	75	..
49	51	53	57	60	61	64	68	71	76	..

20's (N = 85)

47	58	62	64	66	69	70	73	77	80	86
49	58	62	65	66	69	70	73	77	82	88
50	59	62	65	67	69	71	74	77	82	88
51	60	63	65	67	69	72	74	78	82	91
53	60	63	65	67	70	73	75	78	84	94
53	61	63	66	68	70	73	75	78	84	..
57	61	64	66	68	70	73	76	80	84	..
58	62	64	66	69	70	73	76	80	85	..

30's (N = 95)

41	58	61	65	69	71	73	76	79	83	89
45	58	61	65	69	71	73	76	80	83	90
50	58	62	65	69	72	73	77	80	83	90
54	59	62	65	70	72	73	77	81	84	90
55	60	62	66	70	73	74	78	81	85	93
55	60	63	66	70	73	74	79	81	87	..
56	60	65	66	70	73	74	79	82	87	..
56	61	65	67	71	73	74	79	82	87	..
56	61	65	67	71	73	75	79	83	89	..

## APPENDIX "C" (continued)

40's (N = 82)

39	52	57	61	63	65	66	69	73	79	84
42	54	58	61	63	65	67	70	73	81	85
44	56	58	61	64	65	67	70	74	81	..
46	56	58	61	64	65	68	71	75	82	..
47	56	58	61	64	65	68	71	76	83	..
49	57	59	62	64	65	68	72	76	83	..
50	57	60	62	64	65	69	72	77	83	..
51	57	60	62	64	65	69	72	78	84	..

50's (N = 67)

34	48	51	57	60	64	65	69	74	85
41	49	53	58	61	64	65	69	75	86
42	49	54	58	62	64	66	70	76	86
43	49	55	59	63	64	66	70	77	87
45	49	55	59	63	64	67	71	80	..
47	50	55	59	63	65	68	72	80	..
47	51	56	60	64	65	69	74	81	..

60's (N = 57)

34	44	45	49	53	55	62	66	71	81
40	44	46	49	54	57	64	66	72	81
40	44	47	52	55	60	65	67	72	84
41	44	47	52	55	61	65	68	74	..
41	44	48	53	55	61	65	68	74	..
42	45	48	53	55	61	65	69	76	..

70 and over (N = 35)

39	44	47	49	51	55	57	59	71
41	45	47	50	51	56	57	60	74
41	45	47	50	51	56	58	61	79
43	47	49	50	54	56	59	67	..

## APPENDIX "D"

OCCUPATIONS OF SUBJECTS IN A TEST OF  
CRITICAL THINKINGI. Female:

Accountant	Secretary
Bank Teller	Stenographer
Cashier	Teacher:
Clerk (Office)	High School
Clerk (Sales)	Public School
Dietitian	Physical Education
Housewife	Supervisor
Laboratory Assistant	Student:
Locker Room Attendant	Public School
Librarian	High School
Missionary	University
Music Teacher	Other (Nurse,
Nurse	Commercial, etc.)
Public Relations Officer	Waitress

II. Male:

Accountant	Foreman
Analyst	Grocer
Armed Services:	Insurance:
Airforce Officer	Salesman
Army Officer	Underwriter
Auditor	Supervisor
Bank Manager	Immigration Officer
Business Executor	Jeweller
Cabinetmaker	Landlord
Caretaker	Lawyer
Carpenter	Lineman
Chemist:	Maintenance Worker
Biochemist	Machinist
Chemical Engineer	Mechanic
Research Chemist	Meteorologist
Clergyman	Microbiologist
Commercial Traveller	Office Worker:
Civil Servant	Clerk
Construction Worker	Manager
Economist	Photographer
Factory Administrator	Planner (Aircraft)
Farmer	Police Constable

## APPENDIX "D" (continued)

Publisher	Technician
Railway:	Agricultural
Agent	Appliance
Clerk	Chemical
Engineer	Laboratory, Jr.
Supervisor	Laboratory, Sr.
Research Worker	Tool Designer
Salesman:	Transit:
Commercial	Carman
Stocks	Supervisor
Secretary	Veterinarian
Statistician	Welfare Worker
Social Worker	
Stores Keeper	
Teaching:	
Public and High School	
Principal	
Registrar	
Superintendent	
University	

## APPENDIX "E"

TABLE VII  
 "EMOTIONALLY TONED" AND "NEUTRAL" ITEMS  
 OF THE WATSON-GLASER CRITICAL THINKING APPRAISAL

Emotional		Neutral	
Statement No.	Item No.	Statement No.	Item No.
3	11, 12, 13	2	6, 7, 9, 10
6	28	4	17
12	46	5	23
13	49	7	31
15	54, 56	10	37, 38, 39
20	70	11	40, 41, 42
23	78	17	60
26	87, 88	18	62
27	91		
30	98, 99		

B I B L I O G R A P H Y

## BIBLIOGRAPHY

1. Alpern, M. L. "Ability to test hypotheses," Science Education, 1946, 30, 220-229.
2. Anderson A., and Dvorak, B. "Differences between college students and their elders in standards of conduct," J. abnorm. soc. Psychol., 1928-29, 23, 286-292.
3. Anderson, H. A. "Critical thinking through instruction in English," Engl. J., Vol. 36, 1947, 73-80.
4. Berg, F. A. "A simple objective technique for measuring flexibility in thinking," J. gen. Psychol., 1948, 39, 15-22.
5. Bloom, B. S., and Broder, Lois J. "Problem-solving processes of college students: An exploratory investigation," Supp. Educ. Monogr., No. 73, Chicago: University of Chicago Press, 1950.
6. Boley, J. O. "'Intellectual' versus 'physical' peak performance; the age factor," J. of Gerontol., 1946, (I,1), 285-6.
7. Bradford, E. J. G. "Suggestion, reasoning and arithmetic," Forum of Educ., 1925, 3, 3-12.
8. Bromley, D. B. "Some experimental tests of the effect of age on creative intellectual output," J. Gerontol., 1956, (II,1), 74-82.
9. Brownell, J. A. "The influence of training in reading in the social studies on the ability to think critically," Calif. J. Ed. Research, 1953, 4, 28-31.
10. Buck, N. M., and Ojemann, R. H. "The relation between ability in scientific thinking and behaviour in situations involving choice," J. exp. Educ., 1942, 2, 215-219.
11. Burt, C. L. Mental and Scholastic Tests, London: Staples Press, Ltd., 1947.

12. \_\_\_\_\_, "The development of reasoning in school children," J. exp. Pedagog., 1919-20, 5, 68-77.
13. Burton, A., and Joel, W. "Adult norms for the Watson-Glaser tests of critical thinking," J. of Psychol., 1945, 19, 43.
14. Cole, Luella. Psychology of Adolescence. Fourth edition. New York: Rinehart and Co. Inc., 1954.
15. DeBoer, J. J. "Teaching critical reading," Elem. Engl. Rev., 1946, 23, 251-254.
16. Dempster, J. J. "An investigation into the use of estimated factor scores in describing and comparing groups of secondary and senior school boys of eleven plus," Unpublished Master's thesis, University of London, 1944.
17. Deutsche, Jean M. "The development of children's concepts of causal relations," Univer. Minn. Inst. Child Welf. Monogr., No.13, 1937.
18. Doppelt, J. E. "The organization of mental abilities in the age range 13 to 17," Teach. Coll. Contr. Educ. No. 962, 1950.
19. Dressel, P. L., and Mayhew, L. B. Critical Thinking in Social Science. Dubuque, Iowa: Wm. C. Brown Co., 1954.
20. Edwards, T. B. "Measurement of some aspects of critical thinking," J. exp. Educ., 1950, 18, 263-278.
21. Eysenck, M. D. "The psychological aspects of aging and senility," J. Ment. Sci., 1946, 92, 171-181.
22. Fawcett, H. P. "The nature of proof," 13th Yearb. nat. Counc. Teach. Math., 1938, 11-12.
23. Finch, H. R. "How to teach students to read mass magazines critically," Engl. J., 1949, 38, 388-391.
24. Forhan, T. G. "Methods of teaching problem-solving in arithmetic," Cath. Educ. Rev., 1934, 32, 269-282.

25. Foulds, G. A. "Variations in the intellectual activities of adults," Amer. J. of Psychol., 1949, 62, 238-246.
26. Franklin, S. P. "Comprehension of the sayings of Jesus," Univ. of Iowa Studies in Character, 1928, 2, 1-63.
27. Furst, E. J. "Relationship between tests of intelligence and tests of critical thinking and of knowledge," J. Educ. Rev., 1950, 43, 614-625.
28. Gans, Roma. "A study of critical reading comprehension in the intermediate grades," Teach. Coll. Contr. Educ., No. 811, 1940.
29. Gellerman, L. W. "The double alternation problem; II. The behaviour of children and human adults in a double alternation temporal maze," J. genet. Psychol., 1931, 39, 197-276.
30. Glaser, E. M. "An experiment in the development of critical thinking," Contributions to Education, No. 843. New York: Bureau of Publications, Teachers College, Columbia University, 1941.
31. Goodenough, Florence L. Developmental Psychology. Second edition. New York: D. Appleton Co., 1945.
32. Griffiths, R. The Study of Imagination in Early Childhood and its Function in Mental Development. London: Routledge, 1935.
33. Harris, F. E. Jr. "Do children think critically about classroom procedures?" J. educ. Psychol., 1948, 39, 52-59.
34. Harter, Genevieve L. "Overt trial and error in the problem-solving of preschool children," J. genet. Psychol., 1930, 38, 361-372.
35. Hazlitt, Victoria. "Children's thinking," Brit. J. Psychol., 1930, 20, 354-361.
36. Heglin, H. J. "Problem-solving set in different age groups," J. Gerontol., (II, 3), July, 1956.
37. Heidbreder, Edna. "Studying human thinking," in Methods of Psychology, T. G. Andrews, ed. New York: Wiley, 1948.

38. \_\_\_\_\_, "An experimental study of thinking,"  
Arch. Psychol., Vol. II, No. 73, New York, 1924.
39. \_\_\_\_\_, "Problem-solving in children and adults,"  
J. genet. Psychol., 1928, 35, 522-545.
40. Hurlock, Elizabeth B. Developmental Psychology. New  
York: McGraw-Hill Book Co. Inc., 1953.
41. Johnson, H. C. "Problem-solving in arithmetic; a  
review of the literature," Elem. sch. J., 1944, 44,  
396-403, 476-482.
42. Jones, H. E., and Conrad, H. S. "The growth and decline  
of intelligence: a study of a homogeneous group  
between the ages of ten and sixty," Genet. Psychol.  
Monogr., Vol. 13, No. 3, 1933, 223-298.
43. Jones, V. "Disagreement among teachers as to right and  
wrong," Teach. Coll. Rec., 1929, 31, 24-36.
44. Kay, Sylvia C. "Critical reading: Its importance and  
development," Engl. J., 1946, 35, 380-385.
45. Knipp, Helen B. "The development of thinking and of  
concepts," Elem. Engl., 1951, 28, 290-297.
46. Kottmeyer, W. "Classroom activities in critical read-  
ing," Sch. Rev., 1944, 52, 557-564.
47. Lehman, H. C. "Age and achievement," Scientific  
Monthly, 1945, 61, 127.
48. Maier, N. R. F. "Reasoning in children," J. comp.  
Psychol., 1936, 21, 357-366.
49. Miles, C. C., and Miles, W. R. "The correlation of  
intelligence scores and chronological age from  
early to late maturity," Amer. J. Psychol., 1932,  
44, 44-78.
50. Miles, W. R. "Measure of certain human abilities  
throughout the life span," Proc. nat. Acad.  
Science, 1931, 17, 627-633.
51. \_\_\_\_\_, "Age and Human ability," Psychol. Rev., 1933  
40, 99-123.

52. Miller, Jemima, and Weston, Grace. "Slow learners improve in critical thinking," Social Educ., 1949, 13, 315-316.
53. Moore, T. V. The Reasoning Ability of Children in the First Years of School Life. Baltimore: Williams and Wilkins, 1929.
54. Morgan, Antonia B. "Differences in logical reasoning associated with age and education," in Psychological Reports, Vol. II, No. 3. Louisville, Ky.: Southern University Press, 1956.
55. Morgan, J. H. B., and Morton, J. T. "The distortion of syllogistic reasoning produced by personal convictions," J. soc. Psychol., 1944, 20, 39-59.
56. O'Connor, J. The Unique Individual. Boston: Human Engineering Laboratory, 1948.
57. Piaget, J. Judgment and Reasoning in the Child. New York: Harcourt, Brace and Company, 1928.
58. Pingry, R. E. "Critical thinking; what is it?" Math. Teach., 1951, 44, 466-470.
59. Pyle, W. H. "An experimental study in the development of certain aspects of reasoning," J. educ. Psychol., 1935, 26, 546.
60. Raths, L. E. "Thinking test," Educ. Res. Bull., Ohio State Univ., Vol. 23, 1944, 84, 72-75.
61. Raven, J. C. "The standardization of progressive matrices," Brit. J. med. Psychol., 1941, 19, 137-150.
62. Roberts, K. E. "The ability of pre-school children to solve problems in which a simple relationship is kept constant," J. genet. Psychol., 1932, 40, 118-135.
63. Robitaille, H. J. "The development of a test for reasoning ability and its application to students majoring in Science, Classics and Commercial courses," Stud. Psychol. Psychiat. Cathol. Univ. Amer., Vol. 6, No. 3, 1943.
64. Russell, D. H. Children's Thinking. New York: Ginn and Company, 1956.

65. \_\_\_\_\_, "Reading for critical thinking," Calif. J. Elem. Educ., 1945, 14, 79-86.
66. \_\_\_\_\_, "The development of thinking processes," Rev. Educ. Research, 1953, 23, 136-140.
67. Shaffer, L. "Children's interpretation of cartoons," Teachers College Contributions to Education, No. 429, 1930.
68. Shukla, M. M. "A study of inductive reasoning in children of about thirteen years of age." Unpublished Master's thesis, University of London, 1948.
69. Speakman, D. "The effect of age on the incidental re-learning of stamp values; the use of deduction in a subsidiary test," J. of Gerontol., 1954, 9, 162-167.
70. Taba, Hilda. "The evaluation of critical thinking," 13th Yearb. nat. Council soc. stud., 1942, 123-175.
71. Terman, L. M., and Merrill, Maud A. Measuring Intelligence: A Guide to the Administration of the New Revised Stanford-Binet Tests of Intelligence. Boston: Houghton, 1937.
72. Thelen, H. A. "A methodological study of the learning of chemical concepts and of certain abilities to think critically in freshman chemistry," J. exp. Educ., 1944, 13, 53-75.
73. Thorndike, R. L. "Growth of intelligence during adolescence," J. genet. Psychol., 1948, 72, 11-15.
74. Thurstone, L. L. "Primary abilities," Occupations, 1949, 27, 527-529.
75. Watson, G. and Glaser, E. M. Watson-Glaser Critical Thinking Appraisal; Manual of Directions, Yonkers: N. Y.: World Book Co., 1952.
76. Wechsler, D. The Measurement of Adult Intelligence. Baltimore: Williams and Wilkins, 1944.
77. Welford, A. T. Skill and Age, An Experimental Approach. New York: Oxford University Press, 1951.

78. Werner, H., and Kaplan, Edith. "The acquisition of word meanings; a developmental study," Monogr. soc. res. Child Development, 1950, 15.
79. Zant, J. W. "Critical thinking as an aim in mathematics courses for general education," Math. Teach., 1952, 45, 249-256.
80. Zubek, J. P., and Solberg, Patricia. Human Development. New York: McGraw-Hill Book Company, Inc., 1954.