A QUALITATIVE ANALYSIS OF VOCABULARY RESPONSES IN EARLY AND LATE MATURITY

THESIS FOR THE DEGREE OF PH. D.
MICHIGAN STATE UNIVERSITY

JOHN C. EHRMANN

1955

THESIS

This is to certify that the

thesis entitled
A QUALITATIVE ANALYSIS OF
VOCABULARY RESPONSES IN
EARLY AND LATE MATURITY

presented by

JOHN C. EHRMANN

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Psychology

Donald W. Johnson
Major professor

Date July 25, 1955

O-169

N.N.

~~Jan 27, 1978~~ 333
~~Book not stayed~~

A QUALITATIVE ANALYSIS OF VOCABULARY RESPONSES IN EARLY AND LATE MATURITY

by

JOHN C. EHRMANN

A THESIS

Submitted to the School of Advanced Graduate Studies of Michigan State University of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

1955

ABSTRACT

An attempt was made to substantiate the general hypothesis that vocabulary definitions of aged individuals, when analyzed on a qualitative basis, will demonstrate marked intellectual decline. More specifically, it was assumed that aged normal individuals would produce word definitions qualitatively inferior to those produced by young normal individuals.

An experimental group of 100 aged white males (Mean age 75 years) and a control group of 100 young white males (Mean age 25 years) were matched on the basis of intelligence (Wechsler-Bellevue Full Scale IQ) and score correct on the Wechsler-Bellevue Vocabulary Test. All definitions for twenty nouns selected from this vocabulary test were subjected to qualitative analysis consisting of five categories: 1) Class or categorical, 2) Descriptive, 3) Example, 4) Functional and, 5) Error.

The reliability of this scale was tested by having an independent rater classify all definitions of every fourth protocol, fifty in all. Comparison of the ratings indicated 96 per cent agreement.

The following five specific hypotheses were tested:

1) The young group will produce significantly

more Class or categorical definitions

2) The aged group will produce significantly more Descriptive definitions

3) The aged group will produce significantly more Example definitions

4) The aged group will produce significantly more Functional definitions

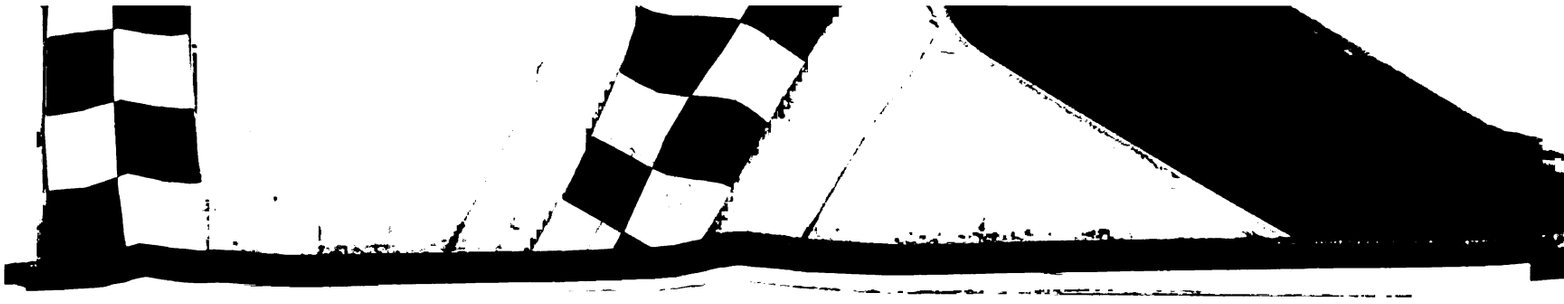
5) The two groups will not demonstrate significant differences in errors of word definitions.

Obtained results substantiated four of the five hypotheses. The young group produced a significantly greater number of Class or categorical definitions. The older group produced significantly greater numbers of Descriptive definitions and Functional definitions. Mean differences between the two groups were significant beyond the 0.5 per cent level of confidence. No significant difference was obtained between the two groups in their total number of failures for the twenty words. Contrary to one of the specific hypotheses, no significant difference was noted for the two groups in the frequency of Example types of word definitions.

In the discussion of the results, the following three aspects were emphasized:

1) The theory that vocabulary represents "old learning" and therefore is resistant to decline in mental impairment is challenged.

2) The validity of psychological tests using a



quantitative measure of vocabulary as an index of the individual's previous level of functioning is challenged.

3) The efficiency of a qualitative analysis of vocabulary word definitions as an indication of the presence and extent of mental deterioration is emphasised.

Approved Donald M. Johnson

Date July 25, 1955

ACKNOWLEDGMENTS

The author's primary debts of gratitude are to Dr. Albert I. Rabin and Dr. Donald M. Johnson.

As the initial chairman of the Guidance Committee, Dr. Rabin provided constructive criticism and patient guidance throughout the formulation of this study.

A particular debt of gratitude is owed to Dr. Johnson, who assumed the responsibility of Guidance Committee Chairman in the absence of Dr. Rabin. He gave freely of his time, encouragement and invaluable advice.

Other committee members who contributed their sage advice and offered constructive criticisms were Dr. Alfred G. Dietze and Dr. Paul Bakan.

The Medical and Surgical Staff of the Veterans Administration hospital at Saginaw, Michigan gave unlimited cooperation to the author and greatly simplified the matter of gathering the data.

This research is dedicated to my wife, Dolores, whose contributions in editing and typing the final manuscript were imeasurable.

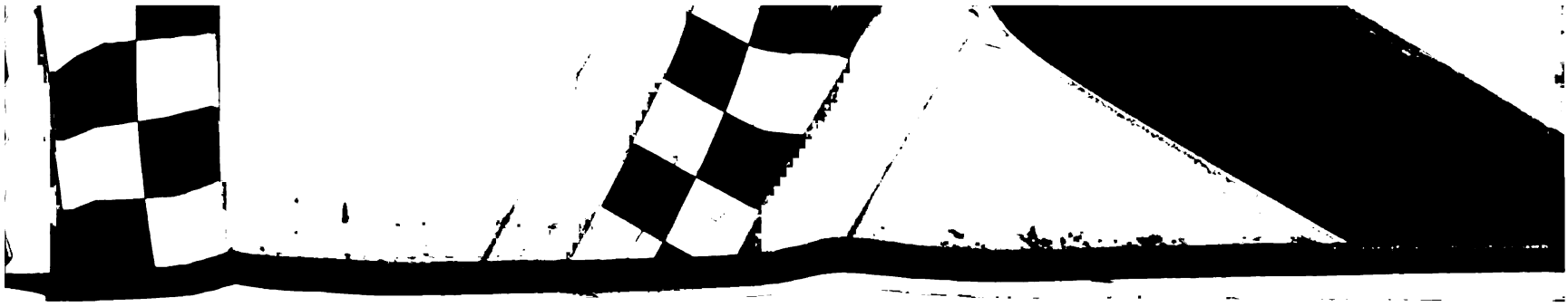
TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

Within recent years a highly significant amount of research has been centered upon the psychological problems encountered in later maturity. It is unfortunate that the realization of the tremendous needs for such research with the aged and aging has been so long in coming, and that these investigations have, as yet, done little more than touch the surface. In comparison with the amount of work that has been devoted to other age groups, the field of gerontology is such a relatively new and unexplored area that the research worker is still in the process of finding and perfecting techniques and methodology applicable to the problems at hand. At the present time psychological explorations of this age range are devoted primarily to a description of the nature of psychological changes in the aging process from the time of maturity up to old age.

Well controlled psychological studies of aged people seem to have been restricted to investigations of abnormal groups including neurotics, psychotics, and various types of physically disabled individuals. Important as these studies are and will continue to be, in consideration of current population trends, it is certainly of the utmost importance for psychology to attempt to thoroughly investi-



gate the normal aging process. Due to the nation's declining birth and death rates, fewer children are being born and more and more people are living into the older age range. Consequently, from decade to decade, a greater proportion of the population of this country falls into the age classification of later maturity.

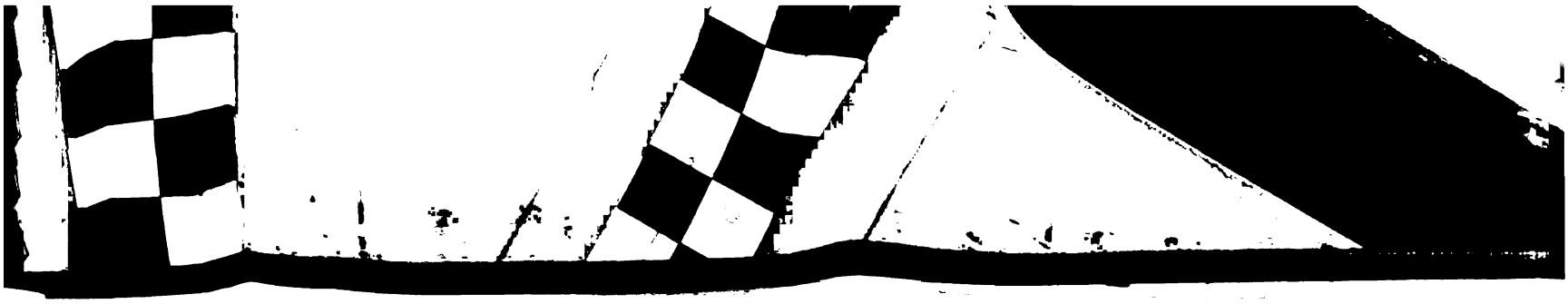
The statistics furnished by the United States Bureau of the Census and the United States Bureau of Labor Statistics provide a dramatic realization of current population trends (82)(83)(84). For example, in 1900, the number of people over sixty-five years of age in this country totaled a little over three million, a ratio of one person of sixty-five or more years to every 25 individuals; by 1950, the Census reported a total of 13.4 million people aged sixty-five or older, a ratio of one oldster to every twelve individuals. Thus the increase in life expectancy over a fifty year period was approximately twenty years, and the population aged sixty-five years or older increased by 37 per cent, whereas the number at ages under sixty-five increased only 13 per cent. Moreover, the Bureau of the Census anticipates still further population gains for older people and estimates that by 1975, close to 21 million people in the United States will be sixty-five years or older, and that by the turn of the century at least 13 per cent of the country's population will have reached the age of sixty-five. Realizing the significance of the above figures, community agencies throughout the country have

insisted that the value to be derived from research devoted to gerontological issues can no longer be denied.

Intellectual Functioning in Later Maturity

Psychology has, from a historical point of view, long been concerned with problems relating to the nature and measurement of intelligence. At the present time, the field of mental measurement has progressed into an advanced and rigorous science, but many problems are still unsolved. Although many valuable instruments have been perfected, it is now widely recognized that it is possible, through the use of these instruments, to obtain only an indication or estimate of actual native intelligence. We are unable to separate any measurement of an individual's functioning from such factors as the motivation to learn and to perform, the general condition and plasticity of the nervous system, and environmental opportunities.

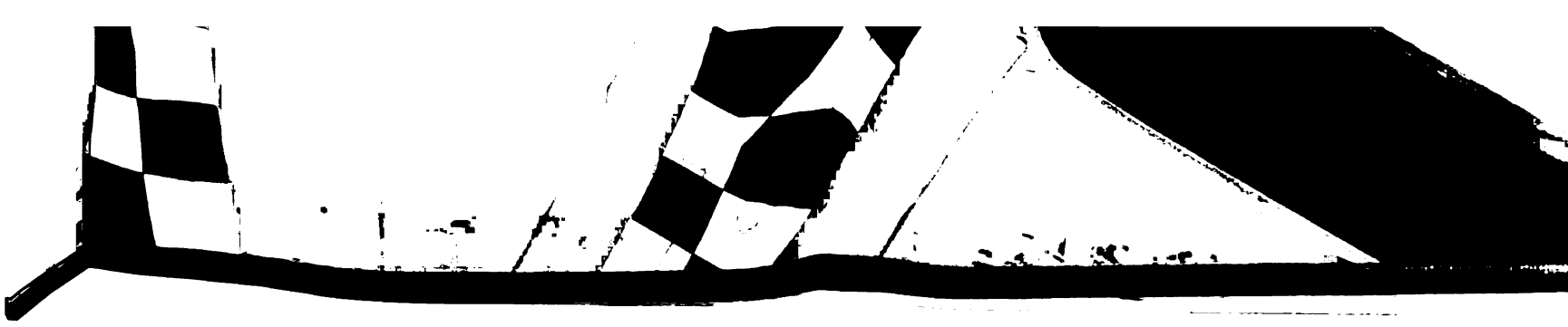
It is generally understood that no matter how carefully and wisely a test of intelligence is constructed, any discussion of its results must be based upon and interpreted in light of the particular manner in which the concept of intelligence is defined and the particular methodology or means utilized in its measurement. In consideration of these facts, the psychologist attempting a comprehensive statement of intellectual functioning during later maturity



is confronted with additional difficulties peculiar to the aged individual.

On the one hand, if an overall or general view of intelligence is adopted, and intelligence is then measured by an overall test score, the gerontologist must be cognizant of the fact that any test or instrument is liable to sample different types of mental abilities from one age level to another. In one study, conducted by Balinski (5), a factor analysis of the Wechsler-Bellevue Scale demonstrated quite different factor loadings from one age range to another. Thus, a test may, perhaps, measure the same mental ability but with varying degrees of reliability and validity for different age groups.

Even if the research worker is provided with some degree of certainty that the same ability can be effectively tapped at various age levels, he is still confronted with serious difficulties when he attempts to interpret his results. Carefully controlled investigations during this period are very few in number and there is an almost complete lack of standardization and normative data to cover the upper age brackets. For example, the 1937 Revision of the Stanford-Binet, though often used to estimate the Intelligence Quotient of adult individuals, reports standardization data only up to eighteen years of age. Even such a widely adopted instrument for the measurement of adult intelligence as the Wechsler-Bellevue Intelligence Scale demonstrates extremely inadequate samples from the upper



age levels. The normative data provided in Wechsler's tables terminates at the age of 59 years and includes a total of less than 200 subjects beyond the age of 50 (86). Since our intelligence tests have primarily been standardized on young people and specifically designed for younger subjects, their use with older individuals is open to the serious challenge that they are not wholly valid in later maturity.

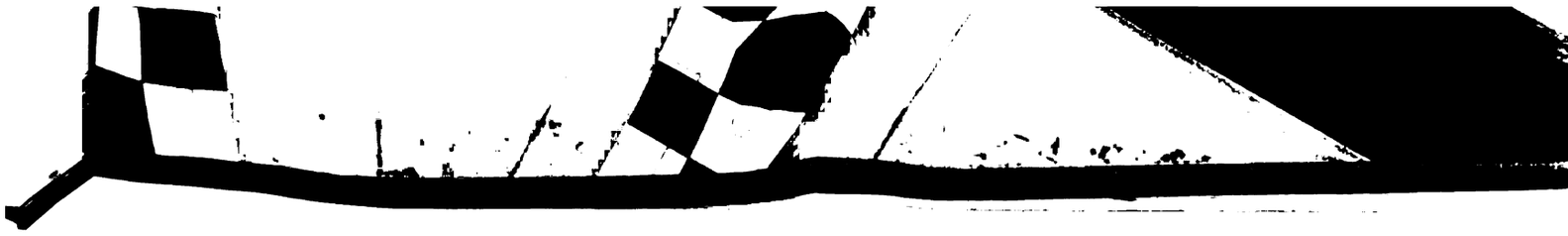
On the other hand, if a more specific definition of intelligence is accepted and measurement is obtained through observation and detailed analysis of specific mental functions, the gerontologist is again faced with problems peculiar to the aged. Judgment, critical thinking, reasoning, and other so-called "higher mental abilities" do not readily lend themselves to quantitative measurement and, thus, are very difficult to compare from one age range to another. In addition, such abilities as reasoning and judgment, when measured on a very narrow basis, tend to be highly correlated with the amount of previous education and training. Thus, in some studies devoted to an analysis of an individual's current functioning level on certain tasks, it is difficult, if not impossible, to estimate the native potential of the aged person because of the many opportunities he has had for continued practice in comparison with a younger subject.

If an attempt is made to adequately match groups of young and old subjects on such variables as education and training, it is necessary to consider the likelihood that the nature of the training of the older individual, which

has usually occurred at a much earlier period, even though similar on a quantitative basis, was probably quite different qualitatively from the more recent training experiences of the younger subject.

The operation of the factor of motivation can readily be seen in the behavior of many aged people. Occasionally, a certain amount of intellectual decline may be directly attributed to the absence of any motivational factors. Since the learning process, with aged individuals, takes a longer amount of time, requires a greater effort, and produces results which are not always as satisfactory as desired, some oldsters cannot or will not readily exert themselves. Thus, general interest wanes and a decrease of intellectual activity and the use of intellectual abilities is apparent. Since motivation does seem to decrease in many aged persons, it is a factor that must always be considered in an investigation of intellectual functioning during later maturity. Because it cannot very easily be quantified, many investigators have chosen to ignore it.

The problem of mental deterioration. It is important to understand exactly what is meant by the concept of "mental deterioration" if one is to objectively study the nature of changes in intellectual functioning that may take place during the aging process. For purposes of diagnostic convenience, mental deterioration is often considered to be one of two rather different types--that which occurs during



later maturity along with the natural increase of age, and that which occurs as a consequence of prolonged mental disease or some brain lesion, and may occur at any age. However, from a psychological point of view there appears to be little difference between these two types. It has frequently been pointed out in the literature that the type of deterioration, evidenced on psychological test performance, encountered in normal old people is highly similar to that encountered in most organic brain diseases and mental disorders (32)(39)(42)(49)(86).

Therefore, at the present time, many people investigating mental deterioration consider this concept as a very general one. For example, Wechsler states, "A person will be considered as giving evidence of mental deterioration when he is no longer able to carry on his intellectual tasks with the speed, accuracy, or efficiency previously characteristic of his functioning level"(86, p.54). A definition such as this one accounts for the fact that mental decline may and does take place independently of any specific organic or mental disease. The decline of mental ability with age should be considered a normal and expected part of the general aging process of the organism as a whole, according to Wechsler and others.

Wechsler feels, "Every human capacity after attaining a maximum begins an immediate decline"(86, p.55). He points to the well established information on changes in the human brain that accompany normal old age--a general loss in mean

brain weight and a gradual shrinking in size--and shows that corresponding changes in intellectual ability closely follow this pattern of deterioration.

The actual task of investigating mental deterioration can be divided into three separate problems:

1. A reliable measure of the individual's actual or present functioning level.

2. The evaluation of the individual's previous functioning level.

3. The expression of the difference between these two measurements in meaningful quantitative terms.

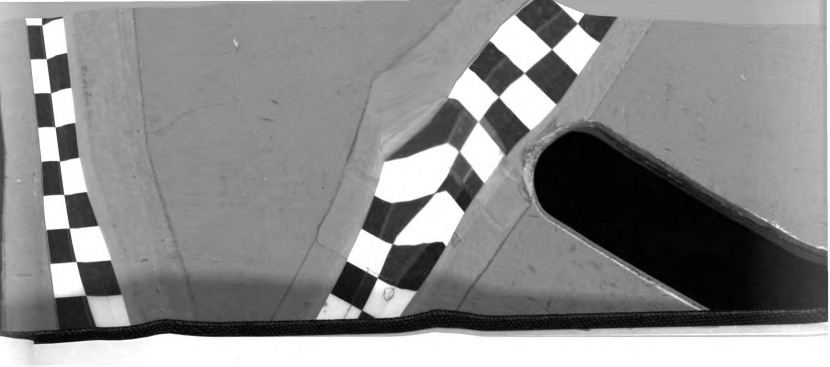
At the present time a fully satisfactory solution has not been found for any of the above problems. The field of mental measurement has, however, established several instruments that purport to reliably estimate the nature of the individual's present functioning. The major difficulties are encountered in the attempt to estimate the previous functioning level of the individual, and thus make a meaningful and quantitative analysis of any possible differences.

The results of intelligence tests. The tremendous growth and expansion of the mental measurement movement during the period of World War I led into psychological surveys of extended age ranges. In one of the earliest studies, Jones and Conrad (43), found a steady intellectual decline after a peak mental level of about 20 years. Their results indicated a decline to a fourteen year level at the age of fifty-

five. Some of the most extensive research in this area was the series of Stanford University Studies on Later Maturity under the supervision of W.R. Miles (60)(61)(62)(63)(64). The results again showed a gradual decline during later years, but found wide individual differences. Beeson (6), found the average mental level of a group of aged individuals to fall within the "dull normal" level. Other studies by Rabin (68), Madonick (57), and Lorge (56), have shown marked intellectual decline in old age.

In general the results of intelligence tests given to many individuals at various periods throughout the life span agree fairly well in showing that after reaching a peak of performance somewhere in the teens or twenties, there is a gradual decline in test scores from early to late maturity. The drop in score between the ages of eighteen and eighty-five years has been found to be as high as 80 per cent of the original score. Wechsler has insisted that all mental abilities reach a maximum usually in the twenties and then begin to decline. He feels that this decline is at first very slow but soon increases perceptibly, and that, once started, it progresses continually (86).

All of these investigations have shown that the amount of intellectual decline varies greatly with the ability being tested. The greatest amount of loss has been found on timed tests of any sort, and problems involving the flexibility of thinking, and the least loss on such tests as vocabulary and general information. There is a marked decline in the speed

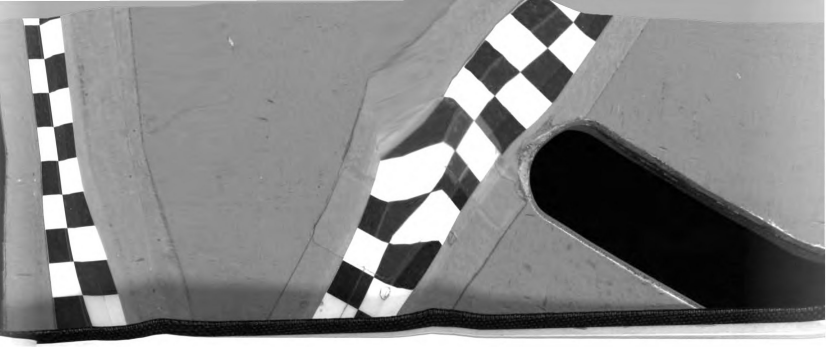


of comprehension of a problem and much difficulty in comprehending material that is entirely new or out of the individual's range of knowledge.

The "Higher Mental Abilities". Scientific studies of the higher mental functions have demonstrated several interesting results. Reasoning and judgment, for example, are among the latest of all mental abilities to reach their peak, and likewise, are among the latest to decline (26)(27)(30)(32)(49)(50). However, when speed and flexibility are included in the measurement of reasoning or judgment, decline is much more marked. The ability to form concepts has been shown to decline greatly in later maturity.

The ability to learn new material and retain items already committed to memory are often listed as common complaints of the aged individual. Investigations have also shown loss of this ability in later maturity (25)(26). One study of rote memory by Kubo (49), found no serious decrement until the age of eighty-two, but several other investigators have shown a marked loss of both verbal and non-verbal memory and learning ability to old age (26)(32)(68)(75). These studies all point out that learning is laborious in the aged person, since he suffers from decreased flexibility of thinking.

There are many examples in the literature of significant creative achievements during later maturity and old age, one indication of high mental level during this period. Dorland



has made several studies of the average age of men in various professions at the time of their most significant creative production, and has found this age to be a good deal older than commonly supposed (18)(19). W.R. Miles, tested imaginative responses to ambiguous stimuli and found no correlation between age in decades and creative imagination (60). Lehman and his associates have reported a series of investigations that showed significant contributions in various fields of science by individuals beyond the age of sixty (52)(53)(54). Although some of the problems of gerontological research mentioned above are avoided when simply considering the relationship between age and intelligence from the basis of the time of occurrence of outstanding examples of the individual's level of intellectual functioning, it is necessary to consider the possibility and even likelihood that the significant achievement of any individual has been maturing and developing over a considerable number of years.

The results of Rorschach tests. Almost all testing of aged individuals with the Rorschach has indicated a decline in the speed of perceiving the ink blot stimuli as meaningful concepts. The amount of material encompassed in any single perceptive response tends to show a steady decline from decade to decade.

Perceptional organization, as measured by the Rorschach test, shows important changes with older individuals. Old people tend to display marked tendencies to make false iden-

tification of objects and show a highly restricted child-like thought content demonstrated by a tendency to use stereotypes (66)(74). Klopfer and others have stated that the Rorschach protocol of the aged individual is much like that of a child and that he tends to define things in terms of the mere utility of the object rather than in terms of its composition or construction (40)(48). In a recent doctoral study, Rochwarg (72) found perceptual development undergoes a decline at senescence. The responses of the aged group seemed to lack appropriate organization and were characteristic of developmentally earlier levels.

The most complete Rorschach analysis of a normal aged group was carried out by Ames et al.(1). These investigators found the aged to be generally slower, less productive, and less efficient than younger persons. They feel that the aged are still able, for the most part, to deal adequately with the practical problems of everyday life, but have a great deal of difficulty in organizing their experiences for future reference. They add that the aged person does not seem to be interested in abstract and theoretical things, or the formulation of original ideas, and indicate that the aged individuals produced many primitive intellectual manifestations similar to those of children. They state, "As the subject deteriorates, his response becomes increasingly like that of a younger and younger person--both qualitatively and quantitatively". On the basis of a qualitative type of analysis of Rorschach factors, they grouped their subjects

into the three classifications of "normal", "pre-senile", and "senile" (1).

The Vocabulary Test

The importance of vocabulary in determining how one person judges another has been heavily stressed recently in popular literature. Within the past few years there has been a steady stream of books dedicated to the aim of teaching an individual to improve his vocabulary. This surge of interest seems to be quite justified in consideration of the fact that a man's vocabulary can yield a great deal of information about himself. The skilled observer picks up such clues about the individual as his educational status, social status, cultural level, and approximate geographical area, all from the manner in which he speaks. The foremost position of vocabulary among all human characteristics by which a person is judged has been pointed out many times in scientific studies (17)(55)(76).

The measurement of vocabulary. It is never possible, of course, to obtain an exact quantification of an individual's vocabulary unless the individual is subjected to the impossible task of defining every known word. As in other areas of mental measurement, general sampling techniques are employed. Usually a given number of words are selected

at random from a standard dictionary with the assumption that the individual's knowledge of these words will provide an index to his total vocabulary. In this fashion such major vocabulary tests as those found in the Stanford-Binet (81) and the Wechsler-Bellevue (86) were devised.

Tests of vocabulary, for convenience, are usually classified in terms of the type of vocabulary they are designed to measure, and in terms of the method of administration. The type of vocabulary tested may be either a general one such as that found in the Stanford-Binet (78), the Wechsler-Bellevue (86), the Columbia Vocabulary Test (12), and the Wide Range Vocabulary Test (3), or specific tests designed to measure a person's vocabulary in specific fields, such as the Knauber Art Vocabulary Test (55), or the Michigan Vocabulary Profile Test (36).

In terms of the administration, essentially two methods are involved with vocabulary tests. One type, usually verbal, presents a stimulus word and asks the subject to respond with a suitable definition. The other type, usually paper and pencil, introduces not only a stimulus word, but also several alternative words from which the subject chooses the one which most appropriately matches the stimulus word (multiple-choice technique).

Factor analysis of the vocabulary response. Some of the investigations that have reported the results of factor analytic studies of the vocabulary test are quite

interesting and provide some insight into what goes into the vocabulary response of an individual. Corter (16), has reported that in his study the following factors were found to have a heavy loading in vocabulary:

a) a learned or academic factor

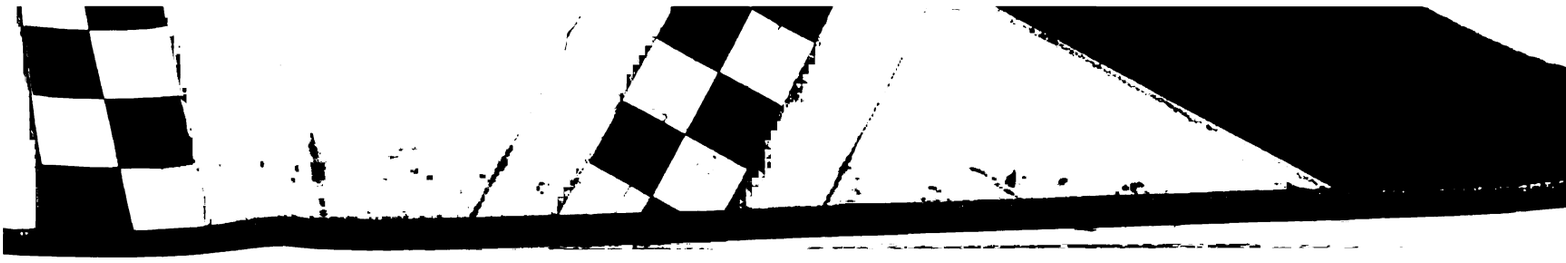
b) a judgment factor involving discrimination

c) a flexibility factor involving the ability to shift one's mental set.

Other similar investigations have shown that the vocabulary response is based upon such factors as verbal ability, experience, and stored information, and have stressed its importance as an indicator of verbal concept formation ability (34)(9).

Vocabulary and general intelligence. There is almost unanimous agreement in the literature that vocabulary tests are highly correlated with indices of general intelligence. Terman states that the vocabulary test had a higher value than any other in his scale and regarded it as the most valuable single subtest (81). Wechsler considered the extent of an individual's vocabulary to be an excellent measure of his general intelligence (86). Armstrong spoke of the vocabulary test as, "a divining rod of the intellect, particularly sensitive to intelligence" (2).

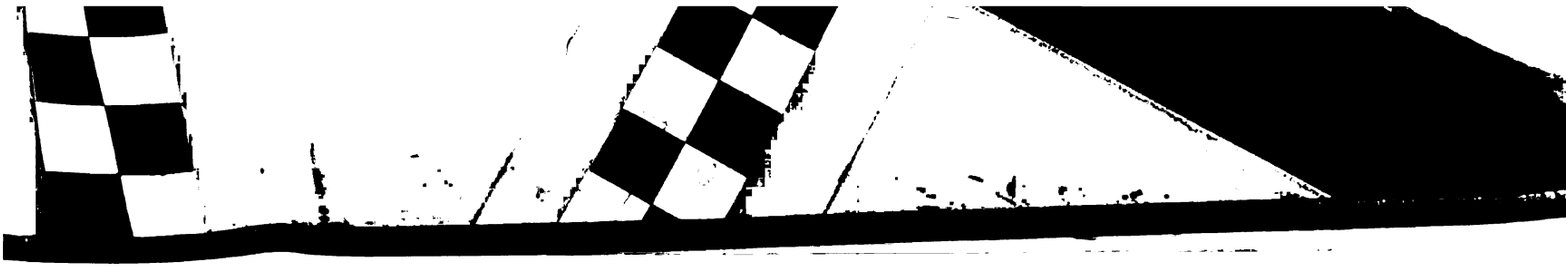
Research reports have substantiated these opinions. Weisenburg, Roe, and McBride (87) reported high correlations between vocabulary score and over-all intelligence scores.



Similarly high positive correlations have been reported by Roe and Shakow (73), Elwood (20), Spache (79), Rabin (67), and Wechsler (86).

Vocabulary and mental deterioration. Research into the nature of intellectual decline in aged and abnormal groups soon established the fact that decline, although quite marked in certain functions, was much less apparent in others. One of the measures of mental ability that continually seemed to show little or no decline with age or mental disease was the vocabulary test. Beeson (6) reporting on his study of intelligence at later maturity found that of all the subtests in his scale, the older people did best by far in the vocabulary sub-test. Green (35), who tested a large group of adults ranging up to 84 years of age, discovered that there was no great change in vocabulary scores after reaching the early twenties. Jones and Conrad (43), working with the Army Alpha, found that vocabulary scores did not show the post-adolescent decline found in almost all of the other sub-tests. Similar results, all indicating that vocabulary test scores held up very well with age and mental disease have been frequently reported (5)(10)(11)(14)(27)(28)(55)(57)(70)(73)(79)(86) and (87). Several investigators such as Rabin (68), Gilbert (30), and Cuff (17), have obtained results indicating that there is a slight tendency for vocabulary scores to actually increase at the ages of later maturity.

Babcock (4) attempted to explain the reason for the



resistance of the vocabulary test to mental impairment and pointed to the results of the physiological findings that cerebral insults primarily damage recently acquired habits. She argued, then, that mental deterioration would be noticed first in new learning and the formation of new habits, and last in the earliest learned material. With the assumption that the vocabulary test represented old habits that had been learned early in life, its resistance to impairment could be accounted for. Starting with this assumption, she utilized a vocabulary test (1916 Terman), as an index of an individual's original level of intellectual functioning, and compared the results of the vocabulary test with the results of efficiency tests designed to measure new learning. She felt that the comparison of these two test scores would produce a valid estimate of the amount of mental deterioration.

More recently, Shipley devised another test measuring mental deterioration in terms of a discrepancy between scores on a vocabulary test and tests measuring an individual's "abstracting" ability (77). This test, the Shipley-Hartford Conceptual Quotient Scale, was originally designed for use with psychotics, but has been used many times to measure deterioration in the aged. Abstraction ability, or the ability to successfully organize one's thoughts into major concepts, has continually been shown to decline in later maturity.

Wechsler has suggested a system or formula to indicate the amount of mental deterioration based upon a comparison

of scores on the sub-tests of the Wechsler-Bellevue Intelligence Scale (86). He has divided his sub-tests into two groups--those which are expected to hold up with age, and those which are not. The vocabulary sub-test, of course, is considered to be one of the best "Hold" scores. The formula devised by Wechsler has been utilized by Rabin (68), Chesrow, Wosika and Reinitz (14), Hunt (41) Magaret and Simpson (58), and many others with quite conflicting results.

Another popular test of mental deterioration, the Hunt-Minnesota Test for Organic Brain Damage, was devised utilizing the 1937 Stanford-Binet Vocabulary Test as a measure of the patient's original mental level and compares this score with the score on tests of the formation and retention of new associations (42). Studies utilizing this instrument have reported contradictory findings (12).

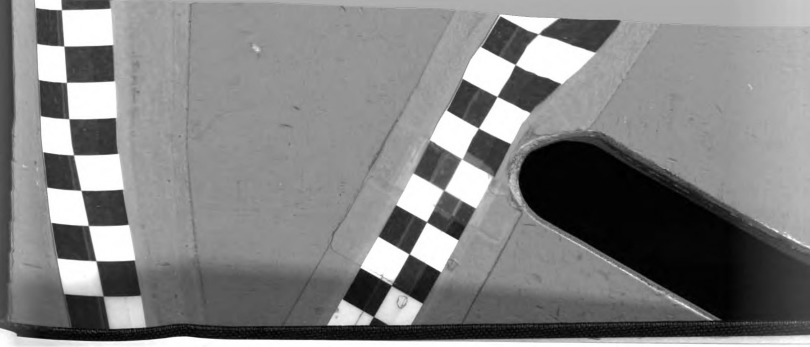
Qualitative analysis of vocabulary. The scoring systems of most vocabulary tests assume that all correct answers are of approximately equal value and do not attempt to differentiate between an excellently verbalized correct answer and a barely acceptable correct answer. The value of accounting for possible differences in the quality of word definitions has been emphasized repeatedly in the literature. Wechsler has stated:

> "Apart from its value as a measure of intelligence, the vocabulary is an especially desirable test to have on any scale because of its qualitative possibilities. In defining a word a subject gives us more than its mere meaning. In many instances he tells us a good deal about himself,

> or at least about the quality and character of his thought processes. These facts may not be immediately apparent, but can be brought to light frequently by analyzing the formal and contextual aspects of the subject's response" (86, p. 99).

As early as 1904, Chambers (13) stressed his opinion that in order to study the development of the use of language in young children, an analysis of the quality and growth of vocabulary was necessary. He tested the vocabulary of a large number of school children and found that he could classify all of the responses into four groups: (a) no answer; (b) wholly wrong answer; (c) vaguely right answer; and (d) correct answer. His results indicated that word definitions of very young children tended to be vague and "a mere outline". Gray and Holmes (34) also noticed that the quality of word definitions changed as children grew older and suggested a qualitative analysis of vocabulary responses.

Binet and Simon (8) reported on the results of their studies of intelligence testing with children. They found that children below the age of nine are prone to define words in terms of their function or usage, and Terman (81) has indicated identical findings. Kirkpatrick (47) discovered, in his work with children and adolescents, that defining a word by categorizing it or giving a synonym for it seemed to represent the highest level of development. He found, too, that "descriptive" types of word definitions did not take place until a late developmental period. Watts (85) has suggested qualitatively grading vocabulary definitions



on a concrete--abstract continuum.

Marx conducted a qualitative analysis of the first fifty words of the 1916 Stanford-Binet vocabulary. He correlated his findings with chronological age and discovered that the highest level of definition was one that categorized the word or gave a synomym for it. Descriptive types of responses were also found to be of high quality, while use, example, and repetition types of responses were considered as inferior.

Green (35) reported the results of her study in which she analyzed the vocabulary responses of large numbers of children and adults on a qualitative basis. She adopted a system of using weighted scores in accordance with the quality of the response and the developmental level of the subject, and found that she could classify all of her data into the following five different levels of quality:

(a) synonym
(b) use and description
(c) explanation
(d) illustration, demonstration, and repetition
(e) error

By grading the responses according to this classification and assigning a differential score for each value of response, a higher correlation between this score and the total test IQ could be obtained than was possible by using just the quantitative vocabulary score.

Feifel and Lorge (23)(24) utilized the same system of

five-fold qualitative analysis and found highly similar results. In their study older children significantly more often produced synonym types of responses, while younger children significantly more often gave inferior types of definitions.

Yacorzynski took note of the qualitative differences possible in a vocabulary response and advanced the hypothesis that the often noted stability of vocabulary test scores in mental deterioration may be more apparent than real; this as a consequence of the fact that the usual score on a vocabulary test is calculated, independently of possible qualitative differences, on an "all-or-none" basis (88)(89). He argued that when responding to a vocabulary test a subject has a wide variety of acceptable responses available to him whereas in many other mental tests there is the possibility of only one correct answer. Certain physiological studies had indicated that many types of brain damage failed to destroy an individual's ability to solve a problem but nevertheless profoundly influenced the method by which the problem was solved, and Yacorzynski felt that it was most reasonable to assume that the mentally deteriorated individual would produce vocabulary responses that were qualitatively much inferior to those of the normal subject, even though both deteriorated and normal individual might receive the same "all-or-none" quantitative test score.

Feifel (22) adopted the Yacorzynski hypothesis and attempted to verify it by testing a total of 370 normals and abnormals (consisting of schizophrenics, manic depres-

sives, and organics), aged 15 to 80 years, matched on the basis of age, education and quantitative vocabulary score. He utilized the five-fold classification schema of Green, which was mentioned earlier, and found that the abnormal subjects at all age ranges selected the inferior types of responses significantly more often than did the normal subjects, these differences being most marked in the responses of older abnormals. Although his samples contained only a small number of cases within the upper age brackets, he found a tendency for older normal individuals to use more inferior responses than young normals, but this difference was not statistically significant. Moran, Moran and Blake compared the vocabulary responses of normals and schizophrenics with this same five-fold schema, utilizing a paper and pencil, multiple choice method. They did not find significant results between the two groups and felt that the differences obtained by Feifel may hold true for organics, but not for schizophrenics (65).

Rabin, King, and Ehrmann (69) conducted a qualitative analysis of responses to the Ammon's Picture Vocabulary Test by groups of normals and schizophrenics matched on the basis of age, sex and education. These investigators stressed the importance of considering the fact that schizophrenia is not a strictly homogeneous group and that individual schizophrenics may vary widely on such a major symptom as "contact with reality". Accordingly, they utilized two groups of schizophrenics--a long term group and a short term group. In

order to account for some of the bizarre verbalizations obtained from their schizophrenic groups, they devised the following qualitative scale:

(a) Synonym or superior definition

(b) Functional or descriptive

(c) Concretistic

(d) Correct, with irrelevant verbalization

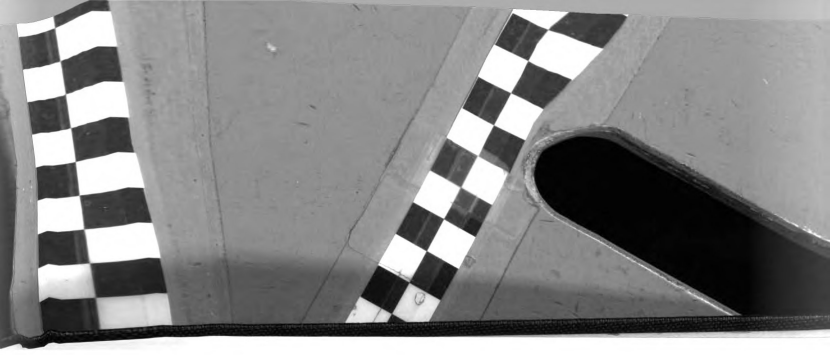
(e) Completely irrelevant

Results using this scale showed significant differences between normals and long-term schizophrenics, but no significant differences between normals and short-term schizophrenics. The authors conclude that with certain types of short-term schizophrenics very little deterioration is evidenced.

Gerstein (29) reviewed the results of a study of concept formation in children by Reichard, Schneider, and Rappaport (71). These investigators discovered that a child utilizes one of three major stages in forming a concept: (1) the concretistic; (2) the functional; and (3) the conceptual. Gerstein pointed out that the same three stages could be utilized to classify definitional levels produced by verbal stimuli and suggested using this as a three-fold system of qualitative analysis for vocabulary responses. She defined this qualitative system as follows:

(a) Concretistic or descriptive--the most primitive type of verbal definition implying memory of an object at a concrete sensory level.

(b) Functional or usage--a more complex method of

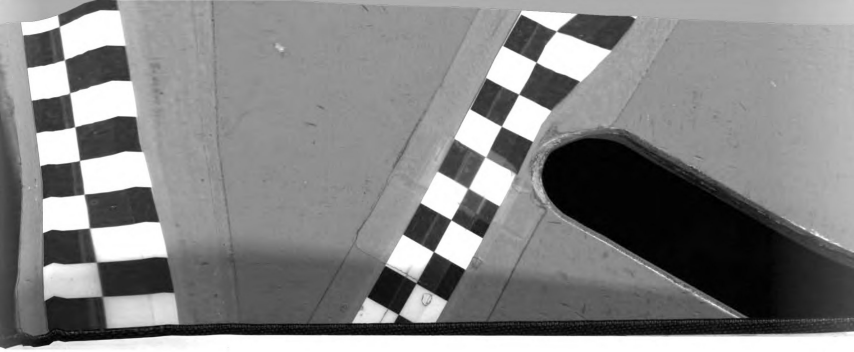


verbalizing a definition--the subject recalls the use to which the object was put in the past.

(c) Categorical or conceptual--the most abstract type of definition.

Gerstein emphasized the relationship between the use of language as manifested by a vocabulary definition and the underlying cognitive processes that may be revealed through a qualitative analysis of these word definitions.

Harrington and Ehrmann (38) utilized both a paper and pencil multiple choice vocabulary test and a verbally presented vocabulary test analyzed according to the three-fold schema suggested by Gerstein and found significant results between normals and schizophrenics on the verbally presented vocabulary scale, but not on the multiple choice, paper and pencil test. These investigators suggest that the complexity of the type of vocabulary response required (simple indication of the correct item from a group of multiple choice answers as compared to producing a verbal definition to an orally presented stimulus word) is important in the funtioning of schizophrenics. Stacey and Portnoy (80) conducted two studies using Gerstein's suggested method of analysis. They tested groups of mentally retarded children and adults and compared them with normals, but found conflicting results. They agree that a qualitative analysis of vocabulary responses is important in attempting to differentiate between varying levels of intelligence but felt that their results indicated that "descriptive" types of word definitions were of a higher



level than were "functional" definitions.

Chordorkoff and Mussen (15) reported on the results of their study in which they utilized a slightly different type of qualitative vocabulary analysis. They felt that it might be profitable to separate the "descriptive" responses from the "functional" responses instead of classifying them at one level as most of the other workers in the area had done. Accordingly they utilized a four-fold schema as follows:

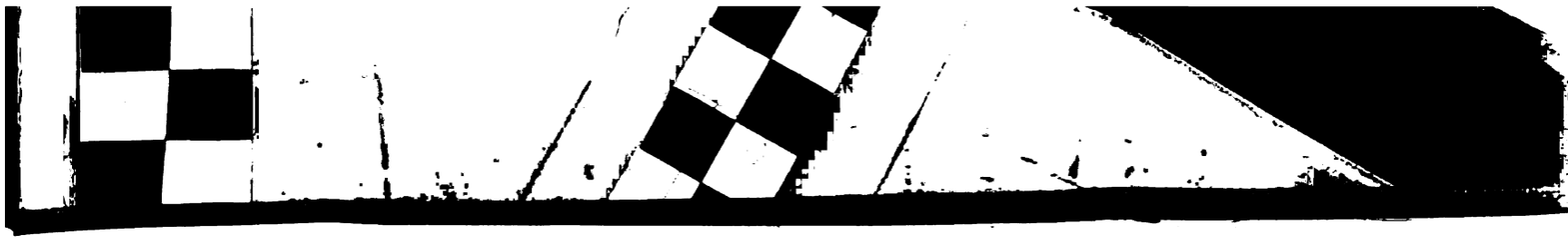
(a) Class or categorical

(b) Descriptive

(c) Example

(d) Functional

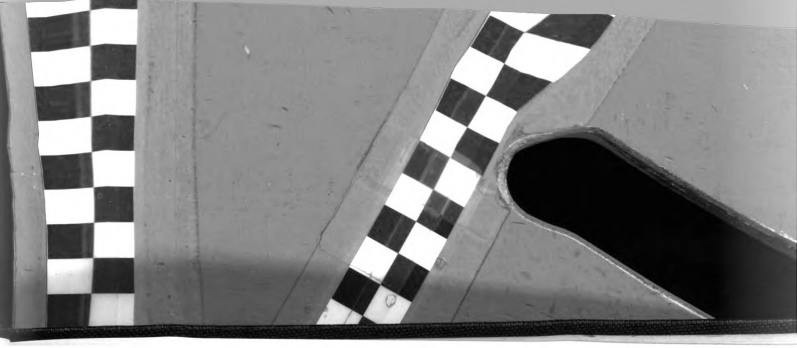
Utilizing a multiple choice version of this scale, they found significant differences between normals and schizophrenics--the normals choosing a greater number of categorical definitions, while the schizophrenics chose a greater number of example and functional responses. No significant differences were noted between the two groups in the frequency with which they chose descriptive definitions, but the normals chose descriptive definitions of high level while the schizophrenics tended to choose descriptive definitions of a low conceptual level.



THE PROBLEM

The performance of deteriorated individuals on vocabulary tests has long been a controversial subject in psychological literature. Many investigators have shown that quantitative vocabulary score declines very little with the advent of old age, mental disorder, or organic pathology, while many other measures of mental functioning show marked impairment. It has been contended by some that the relative lack of impairment on this test provides the clinician with a long needed "index of previous functioning level" from which to judge the degree of mental impairment in other areas. Other investigators have chosen to use the characteristic dispersion of other test scores around the allegedly stable vocabulary score as a device for the differential diagnosis of psychopathology. Several investigators have theorized that this stability of vocabulary test scores is due to the fact that word meanings represent an example of "old learning" and as such is one of the last types of mental functioning to be impaired in mental deterioration.

Other investigators have challenged the stable nature of vocabulary scores and contend that this alleged stability is simply an artifact of the instruments commonly used to measure word meaning. They contend that the reason the vo-



cabulary scores of deteriorated patients show little change is that there are many possible definitions of various levels of quality and, thus, the deteriorated individual is able to earn the same quantitative score as a normal individual, even though his responses are qualitatively much inferior.

Qualitative analysis of vocabulary scores have repeatedly shown significant differences between the performances of normals and deteriorated individuals such as schizophrenics and organics. Some of the studies which have found discrepant results may be criticized on the basis of one or more of the following:

(a) Faulty criteria for selection of experimental groups.

(b) The total number of experimental subjects utilized was inadequate.

(c) Lack of adequate matching of control and experimental groups on such important factors as age, education, and quantitative vocabulary performance.

(d) The techniques utilized to present the vocabulary test to their subjects were not comparable to other studies.

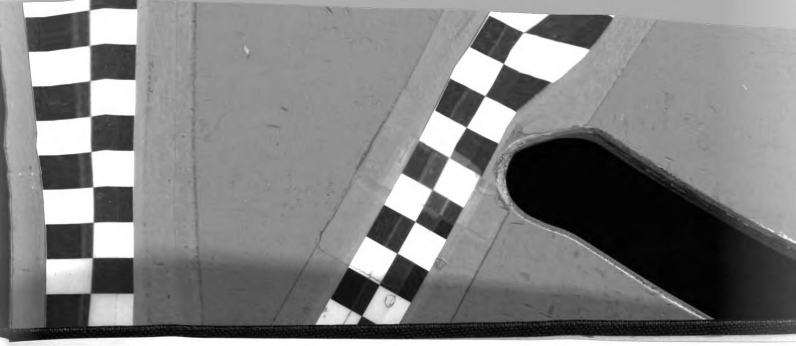
(e) The methods utilized for classification of the vocabulary responses were not comparable to the methods utilized by others.

Studies, mainly utilizing the Rorschach test, have found that perceptual organization undergoes a decline at later maturity. These studies show that the aged individual does not seem to be interested in abstract reasoning and indicate

that his perceptual responses are qualitatively much inferior to those of a normal adult, but very similar to those of children. The conceptual organization of the thought processes of the aged seems to be characteristic of a much earlier developmental level.

Many investigators have pointed out that the mental deterioration that occurs as a normal consequence of old age is not different from that which occurs as a consequence of prolonged mental disease or some brain lesion. Consequently, they advocate defining a deteriorated individual as one who can no longer carry on his intellectual tasks with the speed, accuracy, or efficiency previously characteristic of his functioning level. Gerontologists have established the fact that among the mental functions that decline in later maturity are such abilities as concept formation, discriminative judgment, flexibility of mental functioning, and verbal memory. Factor analytic studies of vocabulary tests have shown that it is precisely these same factors that make up the qualitative aspects of vocabulary responses.

In consideration of the above findings it would seem most reasonable to make the assumption that a carefully controlled investigation of the qualitative aspects of vocabulary responses would demonstrate significant differences between groups of young normal and aged normal individuals. The present investigation proposes an intensive study of this general assumption, utilizing a list of words selected from The Wechsler-Bellevue vocabulary scale. Based on the pre-



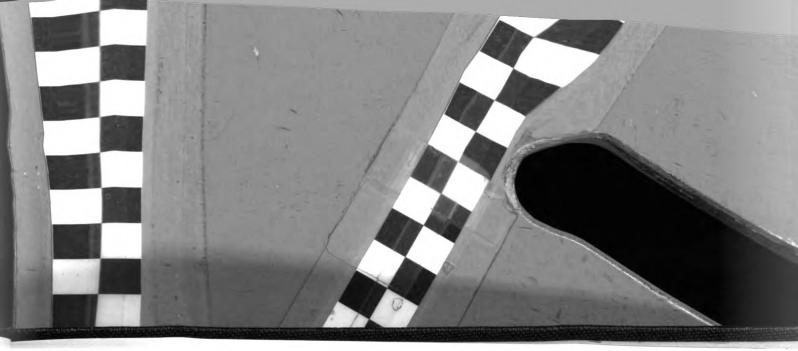
vious discussion and the review of the pertinent literature, the following general hypotheses are offered:

Hypothesis I: There will be significant qualitative differences between the vocabulary performances of young and aged normal individuals.

Hypothesis II: Aged normal individuals will utilize qualitatively inferior vocabulary responses significantly more often than will younger normal individuals.

Subjects

The total population of subjects selected for this study consisted of an experimental group of 100 aged white males, aged 70 to 85, and a control group of 100 young white males, aged 20 to 35. All subjects were patients in the same Veterans Administration general-medical-surgical (GM&S) hospital. All were classified as normal on the basis of having no history of emotional or behavior disorder as far as could be determined, and who in the estimation of the medical and psychological staff did not show any evidence of such a condition while receiving treatment at the hospital. In addition, any patients who presented a history or evidence of degenerative diseases or diseases seriously affecting the central nervous system were not included. The patients were selected from those with ailments that would be least likely to affect performance on any general intellegence test battery.

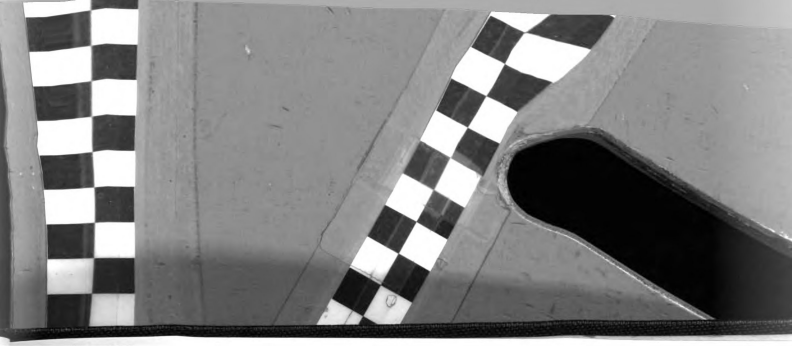


After these initial criteria for selection were met, all of the patients were individually interviewed by the author. The general nature of the study and its purposes were briefly outlined, and a request was made for their participation and cooperation. At this time all patients who did not express willingness to cooperate and displayed no interest in the project were excluded. At this time, too, all patients with a bi-lingual background, or any indication of language handicap were excluded in order to rule out any affects such a background might have on their vocabulary performances.

Matching

At the onset of this study it was hoped that the experimental and control groups could be successfully matched on the basis of sex, intelligence, quantitative vocabulary score, and educational level. Sex was, of course, directly controlled by the selection of the subjects prior to the testing procedure.

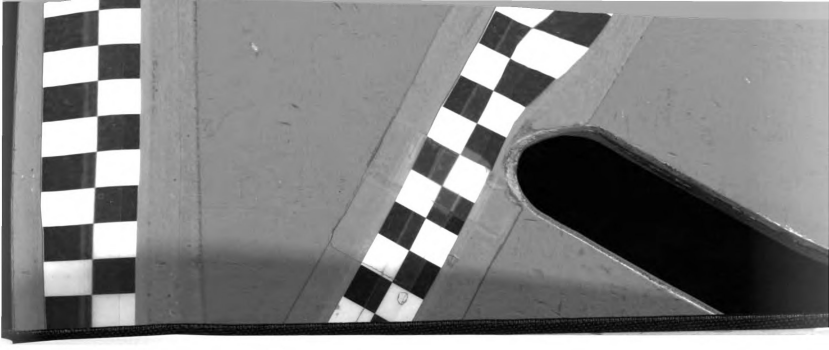
The level of intelligence was derived from the Wechsler-Bellevue Adult Intelligence Scale, Form I. This scale in its entirety was administered individually by the author to each of the subjects in the study. The intellectual level was calculated on the basis of the Full Scale IQ, with the omission of the Vocabulary sub-test, which is normally con-



sidered as an alternative scale. In order to facilitate the final matching of the two groups, and to keep all subjects within practical variations of intelligence, it was decided to utilize only those patients who earned a Full-Scale Wechsler—Bellevue IQ falling into Wechsler's normal classifications (90 to 110). Accordingly, the Wechsler-Bellevue Scale was administered as follows--the Verbal Scale subtests in their usual order of: Information, Comprehension, Digit Span, Arithmetic, and Similarities; this was followed by the Performance Scale subtests in their usual order of: Picture Arrangement, Picture Completion, Block Design, Object Assembly, and Digit Symbol. At this time a Full-Scale IQ was calculated and if this IQ did not fall into the range of 90 to 110, the subject was not included in the study. If his score did fall into the normal ranges of the IQ, the vocabulary sub-test was then administered in the usual way, except that each subject was presented with the entire list of words making up the vocabulary scale, whereas the usual procedure of administration is to continue down the list until 5 successive words have been failed, and then to count any remaining words as automatic failures.

All responses to the vocabulary test were copied in verbatim, and scored according to Wechsler's criteria. The raw scores were converted into Wechsler's weighted scores, in order to simplify matching of the subjects on this variable.

In order to determine the educational level of the



subjects, this study followed earlier investigations and credited either completion of or some attendance at a grade in establishing the number of years of schooling.

Since the subjects for this study were not matched by pairs, mean scores for each of the variables were calculated frequently and selection of succeeding patients was made accordingly.

Table 1 presents all of the relevant information concerning the background of the two groups. It indicates that the mean age of the young group is 25.16 years, the mean age of the aged group is 75.94 years and that the two groups are very well matched on the basis of intelligence, as measured by the Wechsler-Bellevue IQ, and quantitative vocabulary score.

Table 1 also indicates that the two groups did differ significantly on the basis of their educational level. Matching the two groups on this variable simply did not seem to be feasible, for almost invariably the normal aged individual received much less academic training than the young adult of today. When attempts were made to keep the mean educational level of the two groups comparable, it was soon discovered that, by and large, only aged individuals of superior intelligence completed the required amount of schooling. It was decided that a close agreement between the two groups on intelligence test score might, therefore, assure better matching than would close agreement of their academic level.

For the purpose of evaluating the "normalcy" of the

TABLE I

AGE, EDUCATION, IQ, AND VOCABULARY MATCHING DATA FOR THE YOUNG AND AGED GROUPS

Item	Group	
	Young (N=100)	Aged (N=100)
Age		
Range	20-35	70-85
Mean	25.16	75.94
σ	3.64	4.24
IQ (W - B Full Scale)		
Range	90-110	90-110
Mean	102.20	103.36
σ	5.78	5.75
Education (In Years)		
Range	8-16	0-13
Mean	12.44	8.12
σ	1.73	3.45
Vocabulary (W - B Weighted Score)		
Range	7-14	7-14
Mean	10.80	10.86
σ	2.24	2.29

educational level of the two samples utilized in this study, Table II presents a comparison of the mean academic level of the young and aged samples with the median educational achievement of the U. S. population of the same age. This table indicates that the educational level of the two samples in this study is very similar to the educational level of the country at large, showing that the present samples were well selected on the basis of "normalcy".

The Vocabulary Test

Quantitative score. The usual quantitative vocabulary score was computed from the responses given to the 42 items making up the vocabulary sub-test of the Wechsler-Bellevue Intelligence Scale. The vocabulary test from the Wechsler-Bellevue was selected for several reasons. The Wechsler scale had been utilized in order to determine the intellectual level of the subjects, and the administration of its alternative sub-test, vocabulary, was a simple procedure since this sub-test is presented on the usual test blank. Then, too, the reliability and validity of the Wechsler Vocabulary Test as a measure of the subject's "size" of vocabulary or "fund of word knowledge" has been repeatedly demonstrated in the literature. The test has been widely accepted and is, without a doubt, the most frequently utilized of all adult vocabulary scales.

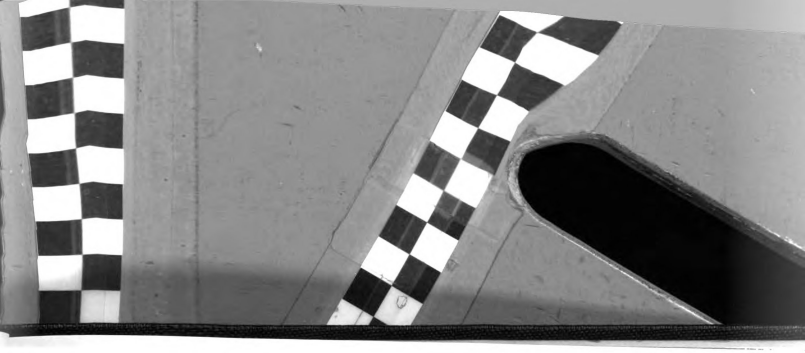
TABLE II

COMPARISON OF THE EDUCATIONAL LEVELS OF THE YOUNG AND AGED GROUPS WITH THE U.S. POPULATION OF THE SAME AGE

	Educational Level
Young	
Mean of present sample, age 20-35	12.44
Median of U.S. population, age 25-29[1]	12.10
Aged	
Mean of present sample, age 70-85	8.12
Median of U.S. population, age 75-79[2]	8.20

[1] Smith, R.M. Educational level of the nations population; School Life, 1953, No. 9, 138-139

[2] Loc. cit.



Qualitative analysis. For purposes of the qualitative analysis itself, twenty words were finally selected. This selection was made from a larger number of words which had originally been compiled from the Wechsler vocabulary scale for a pre-testing study. This larger list, thirty-five words, was administered to approximately 100 individuals during the pre-testing stages. It was soon discovered, however, that many non-nouns did not seem to elicit a wide variety of qualitatively different responses, and so in the final list only nouns were used. In addition, several of the more difficult words toward the end of the list were failed much too frequently and also did not seem to permit a wide variety of responses and were thus eliminated. Some of the words originally selected presented serious scoring difficulties, while others seemed quite ambiguous to many subjects and all of these were likewise eliminated. Finally, then, in this study only the following twenty words were subjected to a qualitative analysis:

1. apple	11. guillotine
2. donkey	12. nitroglycerine
3. diamond	13. stanza
4. nuisance	14. microscope
5. fur	15. belfry
6. cushion	16. affliction
7. bacon	17. pewter
8. nail	18. ballast
9. armory	19. catacomb
10. brim	20. spangle

A detailed analysis of vocabulary responses produced during the pre-testing period indicated that all of the responses could be classified according to the four-fold

qualitative schema advanced by Chodorkoff and Mussen. This schema seemed preferable to many others, for use in this study, since it did not contain categories pertinent primarily to the responses of children and severly disturbed schizophrenics and not frequently encountered in normal adults. This schema also seperately categorizes "Descriptive" and "Functional" types of responses which have been shown to represent wide differences in quality.

Although one of the factors on which both the control and experimental groups were matched was the score correct received on the 42 item Wechsler-Bellevue vocabulary scale, it is evident, since only 20 words selected from this scale were qualitatively analyzed, that the groups could differ widely in terms of the relative number of failures on these specific words. Accordingly, an error category was also utilized in order to compare the relative number of failures for each group. The qualitative schema then, which was used in this study, was five-fold and consisted of the following categories:[1]

(a) Class or categorical responses--the word was defined as a concept and placed into the class or category to which it belonged, or an adequate synonym was produced for it.

(b) Descriptive responses--the word was adequately described in terms of its major properties.

(c) Example responses--an adequate example of the

[1]Sample responses from both groups for all of the words qualitatively analyzed are offered in the appendix.

word was offered.

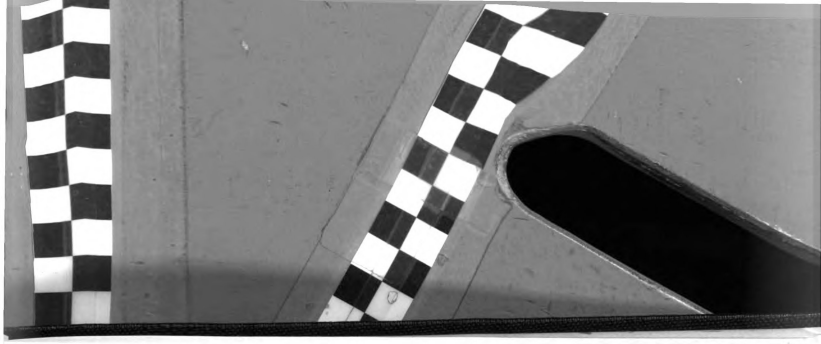
(d) Functional responses--the word was defined in terms of its function or usage.

(e) Error responses--the word was incorrectly defined or no response was produced.

Reliability of scoring. In order to obtain an estimate of the reliability of scoring, every fourth protocol was scored by an independent rater. This independent rater was provided with a complete description of each of the scoring categories and instructions for scoring sample responses. No other instructions or suggestions were given. All identification had previously been removed from the protocols. The scores given by the first examiner were, of course, unknown to the second.

The scores for each protocol were then compared, item for item, to determine the degree of agreement. This agreement was initially computed at 97 per cent. However, a few minor errors in tabulation were encountered. The agreement was again computed, this time at 96 per cent. This high percentage of agreement between two independent scorers indicates that the tests were scored in a reliable manner.

Hypotheses. With reference to the outlined procedure and the general hypotheses stated earlier, a reformulation of these hypotheses in specific operational terms



can now be offered.

A. The young group will produce "Class or categorical" definitions significantly more often than the aged group.

B. The aged group will produce "Descriptive" definitions significantly more often than the young group.

C. The aged group will produce "Example" definitions significantly more often than the young group.

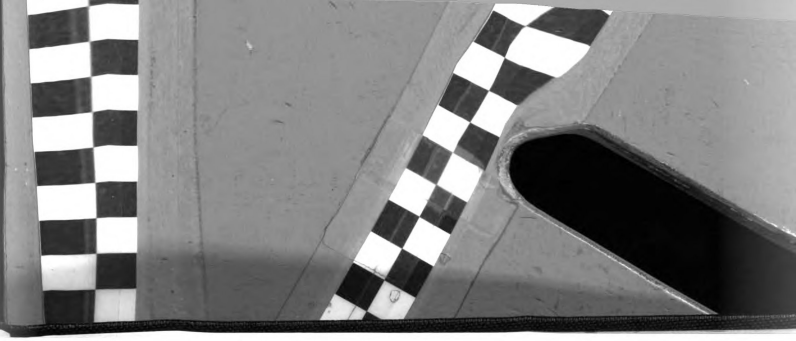
D. The aged group will produce "Functional" definitions significantly more often than the young group.

E. The two groups will not demonstrate significant differences in the frequency of their "Error" responses.

RESULTS

The tenability of five specific hypotheses concerning the qualitative nature of word definitions in an experimental group of aged normal males and a control group of younger normal males was examined in this study. Each of these hypotheses was examined by a t test of the significance of the difference between the means of the two groups. The t test furnished a measure of the validity of the null hypothesis; i.e., a measure of the probability that an obtained result may be attributed to chance. A t level of 2.58 and higher seriously questions the validity of the null hypothesis, and the finding is regarded as highly significant. In each case, since the two groups were not matched pair by pair, the t formula for the difference between uncorrelated means was employed.

A summary of the results obtained is provided in Tables III, IV, V, VI, and VII. In general, all but one of the hypotheses were confirmed by the data. This study indicates that older individuals do manifest qualitative deterioration in conceptualization, as measured by word definitions. The results are examined in detail, in reference to each of the five specific hypotheses.



Hypothesis A. The young group will produce "Class or categorical" definitions significantly more often than the aged group.

Table III presents a comparison between the two groups in terms of the relative number of word definitions which were classified as "Class or categorical". This table indicates that the mean number of "Class or categorical" definitions produced by the young group is 11.30 whereas, the mean for the aged group was only 9.56. The mean difference of 1.74 was significant beyond the 0.5 percent level of confidence and may be interpreted as supporting hypothesis A.

TABLE III

COMPARISON OF THE TWO GROUPS ON THE "CATEGORICAL" TYPE OF RESPONSE

Group	M	σ	σ-Diff	t	Significance
Young	11.30	3.05			
			.415	4.26	Beyond 0.5% Level
Aged	9.56	2.68			

Hypothesis B. The aged group will produce "Descriptive" definitions significantly more often than the young group.

Table IV presents a comparison between the mean number of definitions, produced by each group, which were classified as "Descriptive". This table shows that the mean number of "Descriptive" responses for the aged group is 3.71 and the mean for the young group is 2.95. This mean difference of 0.76 was significant beyond the 0.5 per cent level of confidence and may be interpreted as supporting hypothesis B.

TABLE IV

COMPARISON OF THE TWO GROUPS ON THE "DESCRIPTIVE" TYPE OF RESPONSE

Group	M	σ	σ-Diff	t	Significance
Young	2.95	1.75			
			.287	3.34	Beyond 0.5% Level
Aged	3.71	2.25			

Hypothesis C. The aged group will produce "Example" definitions significantly more often than the young group.

Table V presents the relative number of word definitions produced by each group which were classified as "Example". This table indicates that the mean number of "Example" definitions produced by the aged group was 1.19, while the mean number produced by the young group was 1.14. This mean difference of 0.05 was insignificant. Thus the data of this study, since they demonstrate no significant difference between the two groups in terms of the relative number of "Example" definitions produced by each, do not support hypothesis C.

TABLE V

COMPARISON OF THE TWO GROUPS ON THE "EXAMPLE" TYPE OF RESPONSE

Group	M	σ	σ-Diff	t	Significance
Young	1.19	1.33			
			.191	.26	Insignificant
Aged	1.14	1.37			

Hypothesis D. The aged group will produce "Functional" responses significantly more often than the young group.

Table VI indicates the comparative number of "Functional" responses produced by each group. This table shows that the mean number of "Functional" responses produced by the aged group was 2.41 whereas the mean of the young group was only 1.44. This mean difference of 0.97 was significant beyond the 0.5 per cent level of confidence and may be interpreted as supporting hypothesis D.

TABLE VI

COMPARISON OF THE TWO GROUPS ON THE "FUNCTIONAL" TYPE OF RESPONSE

Group	M	σ	σ-Diff	t	Significance
Young	1.44	1.42			
			.278	3.64	Beyond 0.5% Level
Aged	2.41	2.21			

Hypothesis E. The two groups will not demonstrate significant differences in the frequency with which they produce "Error" responses.

Table VII presents a comparison of the two groups on the frequency of "Error" responses. This table indicates that the mean number of "Error" responses produced by the young group was 3.11 while the mean of the aged group was 3.18. The mean difference of 0.07 was insignificant. Thus, since the data do not indicate any statistically significant differences between the two groups, the results may be interpreted as supporting hypothesis E.

TABLE VII

COMPARISON OF THE TWO GROUPS ON THE "ERROR" TYPE OF RESPONSE

Group	M	σ	σ-Diff	t	Significance
Young	3.11	2.06			
			.215	.22	Insignificant
Aged	3.18	2.51			

In order to present a more graphic summary of the relative number of word definitions classified under each of the qualitative categories, a column graph was drawn showing the results for both the young and aged groups. This is presented in Figure 1.

This figure indicates that for the twenty word vocabulary list the young group produced the following:

56.50 per cent "Categorical" definitions
14.75 per cent "Descriptive" definitions
5.95 per cent "Example" definitions
7.20 per cent "Functional" definitions
15.60 per cent "Error" definitions

This figure also indicates that for the twenty word vocabulary list the aged group produced the following:

47.80 per cent "Categorical" definitions
18.55 per cent "Descriptive" definitions
5.70 per cent "Example" definitions
12.05 per cent "Functional" definitions
15.90 per cent "Error" definitions

FIG. 1. PER CENT OF RESPONSES OF BOTH GROUPS FALLING INTO EACH QUALITATIVE VOCABULARY CLASSIFICATION

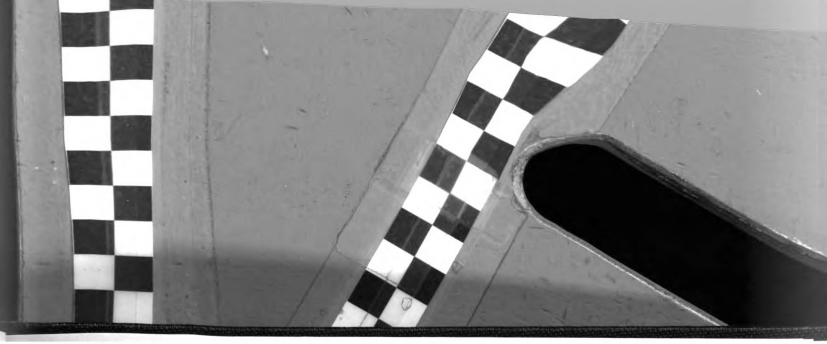
DISCUSSION AND CONCLUSIONS

The basic purpose of this study was to investigate the hypothesis that aged normal individuals do show a decline in their functioning ability on a vocabulary test when their word definitions are qualitatively analyzed. This assumption was based on the reasoning that a quantitative scoring system does not account for the quality of the definitions produced and assumes all correct scores to be of equal value. The results of this study have born out this assumption. The data show that aged individuals produced significantly fewer very superior word definitions of a class or categorical type than younger individuals. The aged group produced significantly more inferior definitions of a descriptive and functional type.

In light of these results, this study lends direct support to those gerontologists who claim that all aspects of intelligence decline in later maturity. These people theorize that mental decline occurs in all areas of mental functioning as a part of the normal aging process. Contrary theories have held that although overall intellectual functioning does decline with age, certain specific aspects of mental functioning, such as vocabulary, show little or no decrement as age increases.

The amount of intellectual decline actually experienced by the aged group utilized in this study is, of course, unknown. The mean Wechsler-Bellevue Intelligence Scale IQ of the aged group was not significantly different from the mean IQ of the younger group, but it is necessary to consider that this does not mean that the actual intelligence test performances of the two groups were the same. The aged group earned their IQ's on the basis of the scores for their test performance plus additional credit for their advanced age. Additional age credit is automatically computed by the use of Wechsler's tables, but these tables continue only through the age of fifty-nine. The usual procedure advocated by Wechsler, for obtaining Wechsler-Bellevue IQ's with individuals aged sixty or older is to multiply the difference between sixty and the subject's actual age by approximately 0.3 and then add this figure to the IQ, as indicated in his tables, that the aged individual would have earned at age sixty. Various studies in the literature have shown that for aged subjects with normal or below normal intelligence this procedure seems fairly accurate, but that the aged individual with above average intelligence attains additional overly generous credit for his advanced age.

At any rate, the actual level of intelligence test performance of the aged group in this study was markedly lower than that of the younger group. Although the pre-deterioration IQ of the aged group, therefore, was probably quite different from the present IQ of the young group of this study, there



is little doubt, due to the accuracy with which the two groups were matched, that they did not significantly differ in the factor of general vocabulary performance or overall word knowledge as measured by quantitative vocabulary test score. These scores were computed, of course, on actual test performance with no additional credit or correction being applied for age. In addition, since both groups produced the same mean number of errors in defining the list of twenty specific words utilized for qualitative analysis, it seems certain that the quantitative vocabulary level of the two groups was the same.

The findings of this study seriously challenge the basic theories underlying current psychological tests of mental deterioration--namely, that vocabulary scores do not decline in later maturity because vocabulary is an index or measure of "old habits" and "early learning", in contrast to other abilities which show marked decline in later maturity and are theoretically assumed to involve more recent association. The findings, rather, indicate that the stability of vocabulary test scores in mental deterioration is an artifact of the usual scoring procedure. They support the Yacorzynski hypothesis that the aged individual can give a correct but inferior and "easier" response and thus not demonstrate his lowered efficiency; whereas a younger and non-deteriorated individual, with superior mental efficiency, will produce an equally correct, but more difficult and highly complex response.

Further interpretation of the results of this study

leads to the conclusion that a vocabulary test score gives an extremely imperfect indication of an individual's early or pre-deterioration level of functioning and that the test instruments commonly utilized to measure mental deterioration possess extremely doubtful validity. It would appear then, that there is little justification for future investigations of mental deterioration in which a quantitative vocabulary test score is interpreted as an index of an earlier level of mental efficiency. It would seem much more justifiable, in light of current knowledge, to utilize a vocabulary test, qualitatively analyzed, as providing some indication of the actual loss of mental efficiency experienced by an individual, since his patterns of response in defining words have been shown to be related to the impairment of mental functioning.

Certainly the manner of approach and the method of procedure used by an individual in defining a word are important in understanding his mental processes. Although a study of the qualitative level of his word definitions will indicate the presence and, to some extent, the degree of mental impairment, this indication will be very tentative and approximate unless a highly refined qualitative analysis is constructed. The five-fold schema utilized in this study can be considered roughly, to represent five different levels of mental ability ranging from superior to inferior. However, much caution must be applied in so interpreting the qualitative scale with results obtained from any single individual. Within each of the four correct categories a wide variation of qual-

ity of response is still quite possible. A class or categorical type of definition, which in general is quite superior to other types of definitions, can range from a simple synonym for the stimulus word to a highly complex genus type of definition in which the stimulus word is modified and precisely specified. A descriptive type of definition can range from an exact specification of the stimulus word in terms of its major properties, to a much inferior and limited description of the stimulus word in terms of certain of its very concrete and relatively minor properties. It is necessary to consider that the level of quality of a word definition might possibly overlap somewhat for some of these qualitative categories. Actually, under special circumstances and in limited conditions, it is quite conceivable that a descriptive or functional type of definition would be superior to the more abstract and complex class type of definition. However, when considering large groups of individuals who are asked to verbally define a word, the qualitative categories of 1) Class or categorical, 2) Descriptive, 3) Example, 4) Function and 5) Error, may be considered to represent a scale of decreasing value of definition. The specific categories, however, do not represent equal intervals of value along the scale.

In normal usage certain words tend to call forth specific types of definitions. Nothing is known, at present, about the exact frequency of the various qualitative types of word definitions the normal individual will produce.

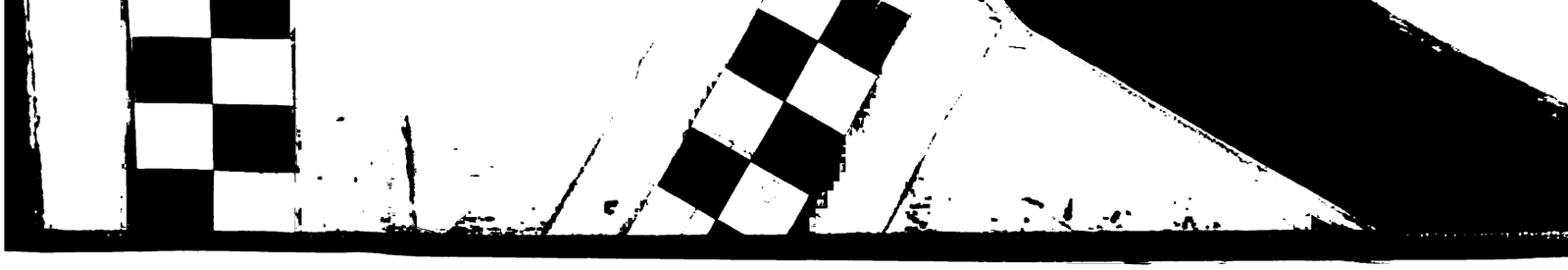
Most frequently, of course, for the normal individual, are the superior definitions such as Class or categorical definitions. However, the more inferior types of word definitions will vary in terms of the frequency with which they are produced, depending upon the specific composition of the vocabulary list utilized. Since the vocabulary list utilized in this study consisted entirely of nouns that had been previously found to elicit a variety of different types of qualitative responses, no real conclusions pertinent to this question can be drawn from the data. Nevertheless, the relative percentages with which each category was elicited are of some interest. In this study both groups produced mainly superior definitions of the class or categorical type, with the younger group having a markedly greater percentage of these responses. Both groups produced a much smaller proportion of inferior definitions of the descriptive and functional types, with the aged group eliciting significantly higher percentages of these. The percentage of functional definitions for the younger group was quite small. Both groups produced the same percentage of example types of definitions, but these definitions were very infrequent and proportionately represented the smallest of all categories for both groups.

The relative proportions of descriptive and functional types of responses elicited by the aged group in this study may possibly have some relation to the Rorschach investigations which have indicated that aged individuals tend to perceive things in a very practical and concrete light, and conceptu-

alize stimuli in terms of their usage. However, in light of the current knowledge pertaining to the nature of decline in mental efficiency during later maturity, this tendency towards descriptive and functional types of responses should be interpreted as signifying a decline in the ability to successfully cope with abstract and complex conceptualizations rather than indicating a specific kind of orientation or mental set of the aged individual toward his environment.

A final word of caution toward interpretation of the findings of this study is offered. The aged individuals tested in this investigation developed and matured in the cultural milieu existent approximately fifty year ago. Just how much this cultural environment differed from that of the younger group is impossible to estimate. The matching data showed, as one example, that educational levels for the two groups were vastly different. There is a possibility that differences in the qualitative vocabulary definitions of the two groups are, to some extent, a reflection of the many cultural changes that took place during this fifty year period.

In addition, the aged people have, for all practical purposes, been away from the formal academic environment for a considerable number of years, while a large proportion of the younger group have only recently completed their schooling. It is certainly possible, therefore, that the younger group in this study would be somewhat more prone or accustomed to offering verbal definitions of a higher conceptual conceptual level.



SUMMARY

An attempt was made to substantiate the general hypothesis that vocabulary definitions of aged individuals, when analyzed on a qualitative basis, will demonstrate marked intellectual decline. More specifically, it was assumed that aged normal individuals would produce word definitions qualitatively inferior to those produced by young normal individuals.

An experimental group of 100 aged white males (Mean age 75 years) and a control group of 100 young white males (Mean age 25 years) were matched on the basis of intelligence (Wechsler-Bellevue Full Scale IQ) and score correct on the Wechsler-Bellevue Vocabulary Test. All definitions for twenty nouns selected from this vocabulary test were subjected to a qualitative analysis consisting of five categories: 1) Class or categorical, 2) Descriptive, 3) Example, 4) Functional and, 5) Error.

The reliability of this scale was tested by having an independent rater classify all definitions of every fourth protocol, fifty in all. Comparison of the ratings indicated 96 per cent agreement.

The following five specific hypotheses were tested:

1) The young group will produce significantly

more Class or categorical definitions

2) The aged group will produce significantly more Descriptive definitions

3) The aged group will produce significantly more Example definitions

4) The aged group will produce significantly more Functional definitions

5) The two groups will not demonstrate significant differences in errors of word definitions.

Obtained results substantiated four of the five hypotheses. The young group produced a significantly greater number of Class or categorical definitions. The older group produced significantly greater numbers of Descriptive definitions and Functional definitions. Mean differences between the two groups were significant beyond the 0.5 per cent level of confidence. No significant difference was obtained between the two groups in their total number of failures for the twenty words. Contrary to one of the specific hypotheses, no significant difference was noted for the two groups in the frequency of Example types of word definitions.

In the discussion of the results, the following three aspects were emphasized:

1) The theory that vocabulary represents "old learning" and therefore is resistant to decline in mental impairment is challenged.

2) The validity of psychological tests using a

quantitative measure of vocabulary as an index of the individual's previous level of functioning is challenged.

3) The efficiency of a qualitative analysis of vocabulary word definitions as an indication of the presence and extent of mental deterioration is emphasized.

APPENDIX

Sample Responses, As Qualitatively Categorized

For The Twenty Vocabulary Words

1. Apple

"Class" responses:

a fruit
an edible pome

"Descriptive" responses:

red and sweet
a food

"Example" responses:

crab apple
Northern Spy

"Functional" responses:

you eat it
make pies from them

2. Donkey

"Class" responses:

an animal
a beast of burden
a jackass

"Descriptive" responses:

it has long ears and four legs
a thing like a horse

"Example" responses:

No "Example" responses were elicited
for this stimulus word.

"Functional" responses:

you ride on it
it pulls wagons and carries things

3. Diamond

"Class" responses:

a gem
a precious stone
a mineral of crystallized carbon

"Descriptive" responses:

it's hard carbon
a polished stone

"Example" responses:

a diamond in a ring
a baseball diamond

"Functional" responses:

for jewelry
used to cut glass

4. Nuisance

"Class" responses:

a pest
a bother
that which annoys or vexes

"Descriptive" responses:

a guy who bothers you
a guy who gets in your hair

"Example" responses:

the guy who kept asking me questions yesterday
my room mate

"Functional" responses:

he holds you up in your work
he gets you mad

5. Fur

"Class" responses:

a pelt
a hide
an animal's skin

"Descriptive" responses:

hairy stuff
an animal's outside parts
part of an animal

"Example" responses:

fox fur
mink fur

"Functional" responses:

make coats out of it
to keep animals warm

6. Cushion

"Class" responses:

a pillow
a pad
a device to deaden jar or impact

"Descriptive" responses:

soft and comfortable thing
a soft seat

"Example" responses:

a car cushion
a pin cushion
cushion on couch

"Functional" responses:

you sit on it
it makes things soft

7. Bacon

"Class" responses:

meat of a pig
pork
dried and salted flesh of a hog

"Descriptive" responses:

meat
a food
narrow strips with eggs

"Example" responses:

No "Example" responses were elicited from this stimulus word.

"Functional" responses:

you eat it
you cook it with eggs

8. Nail

"Class" responses:

a spike
a brad
a slender piece of metal for clenching wood together

"Descriptive" responses:

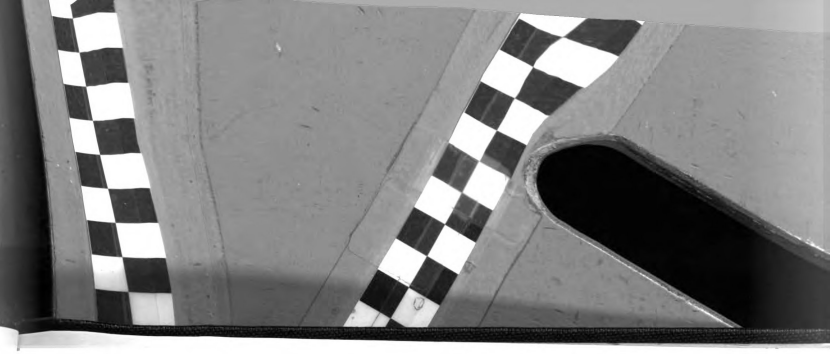
a kind of peg
hard skin
a connecting device

"Example" responses:

ten penny nail
a fingernail

"Functional" responses:

you pound it
to join wood



9. Armory

"Class" responses:

a military building for training soldiers
a building for storing arms
building utilized by the militia for training troops and storing weapons

"Descriptive" responses:

a soldier's club
an army building
National Guard place

"Example" responses:

the big building out on Mill Street

"Functional" responses:

where soldiers practice
to put guns in

10. Brim

"Class" responses:

rim
edge
margin
border

"Descriptive" responses:

the outside part
the top of something

"Example" responses:

a hat brim
brim of a glass

"Functional" responses:

makes the end stronger
keeps sun and rain from eyes

11. Guillotine

"Class" responses:

a means of executing
a way to behead

"Descriptive" responses:

a large weighted knife
an execution blade
the French way of killing people

"Example" responses:

No "Example" responses were elicited for this stimulus word.

"Functional" responses:

to kill people
it chops off heads

12. Nitroglycerine

"Class" responses:

an explosive compound
a medical chemical

"Descriptive" responses:

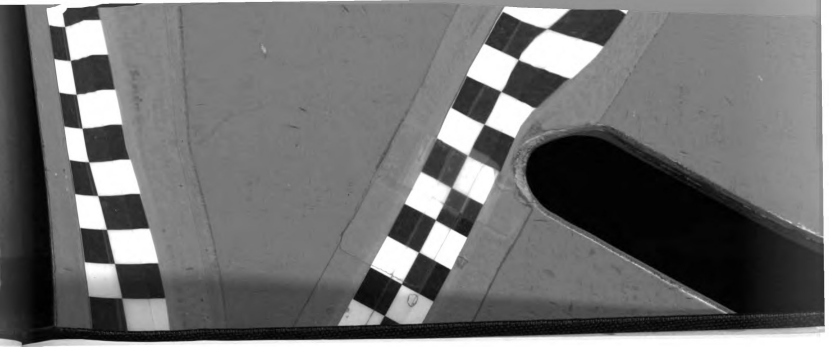
a yellow liquid substance
a combination of glycerol and nitrogen

"Example" responses:

dynamite

"Functional" responses:

to blow things up
it's used in ammunition



13. Stanza

"Class" responses:

a certain number of lines in a poem
a paragraph in poetry or a song
a division of poetry

"Descriptive" responses:

it's part of a poem
it's a phrase

"Example" responses:

No "Example" responses were elicited
for this stimulus word.

"Functional" responses:

to divide
to set parts away from each other

14. Microscope

"Class" responses:

a magnifying instrument
an instrument to magnify or enlarge small objects

"Descriptive" responses:

a black instrument with several lens
a round metal tube with glass prisms inside

"Example" responses:

electron microscope
thing on the table in Dr. Porter's office

"Functional" responses:

to see small objects
to find germs

15. Belfry

"Class" responses:

a church bell tower
a steeple for housing a bell
a place for keeping a bell

"Descriptive" responses:

a tower
top part of a church

"Example" responses:

No "Example" responses were elicited for this stimulus word.

"Functional" responses:

for holding bell
to keep bell in

16. Affliction

"Class" responses:

misfortune
grief
a malady
any distress of body or mind

"Descriptive" responses:

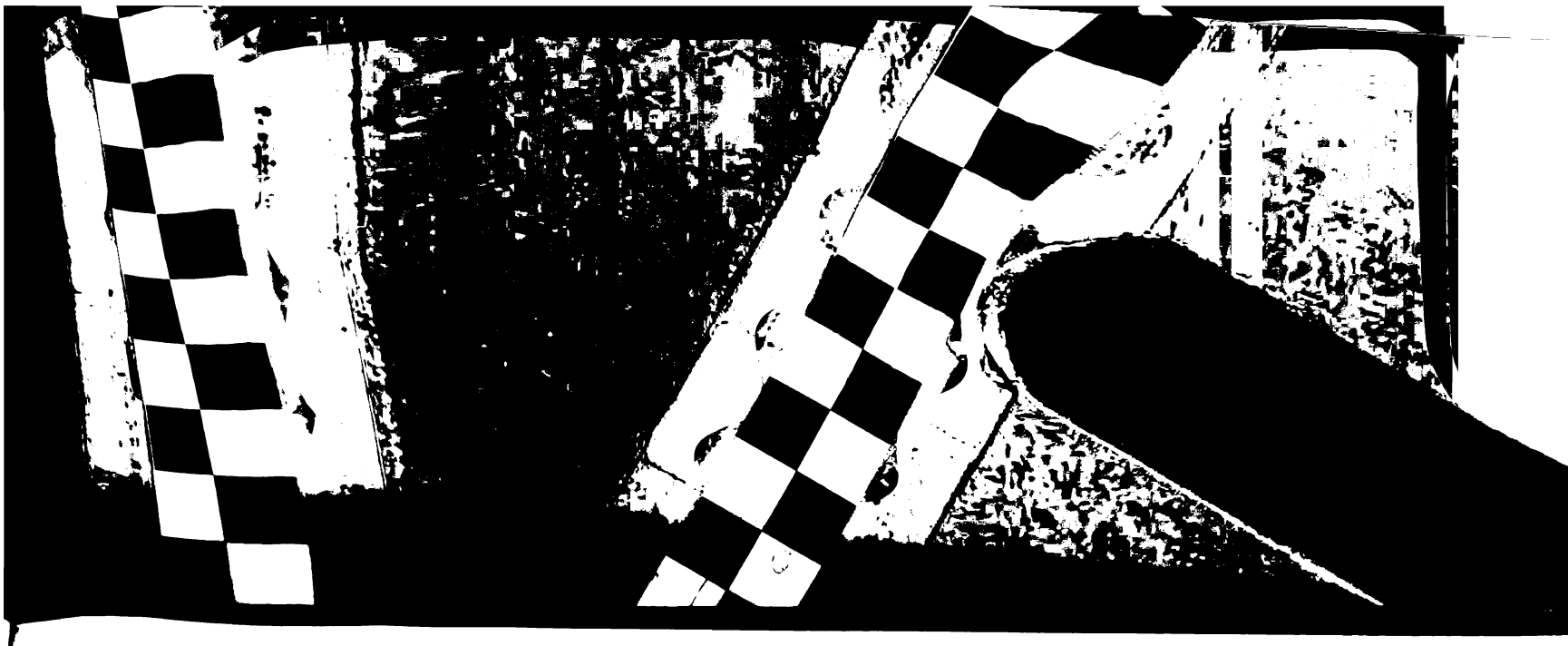
a disease
an unfortunate happening

"Example" responses:

rheumatism
arthritis
a broken leg

"Functional" responses:

something which hurts
it inconveniences you
it makes you sickly



17. Pewter

"Class" responses:

an alloy of lead and tin
a metal of lead, formerly utilized to make dinnerware

"Descriptive" responses:

a metal
it's dark and silver-like

"Example" responses:

the dishes in museums

"Functional" responses:

to make dishes
it's for pottery

18. Ballast

"Class" responses:

a weight used to steady ships
material used to maintain stability

"Descriptive" responses:

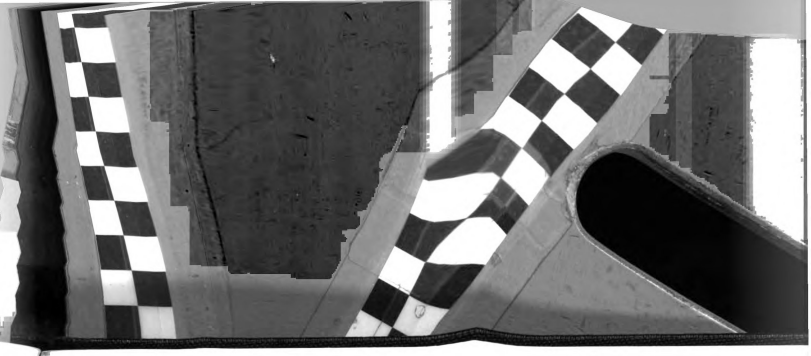
an added weight
a heavy substance

"Example" responses:

lots of sand
filled water tanks
gravel or stones

"Functional" responses:

to balance something
for keeping things even



19. Catacomb

"Class" responses:

an underground cell used by Christians
an ancient burial tomb

"Descriptive" responses:

a cave
an underground tunnel

"Example" responses:

No "Example" responses were elicited
for this stimulus word.

"Functional" responses:

to crawl in and hide
for burying people

20. Spangle

"Class" responses:

a glittering substance used for decoration
shiny things to decorate a costume
bits of sparkling material used to dress things up

"Descriptive" responses:

shiny things
gaudy trimmings
tiny sparkling parts

"Example" responses:

tinsel on a christmas tree
stuff on circus costumes

"Functional" responses:

to make things shiny
for a fancy dress

BIBLIOGRAPHY

1. Ames, L.F., Learned, J., Metraux, R. and Walker, R.M. Rorschach responses in old age. N.Y., Paul Hoeber, 1954.

2. Armstrong, C.P. Tests as diagnostic aids in clinical practice. Amer. J. Ment. Def., 1943, 47, 270-276.

3. Atwell, C.R. and Wells, F.L. Wide range multiple choice vocabulary tests. J. Appl. Psychol., 1937, 21, 550-555.

4. Babcock, H. An experiment in the measurement of mental deterioration. Arch. Psychol., N.Y., 1930, 18, No. 117.

5. Balinski, B. An analysis of the mental factors of various age groups from six to sixty. Genet. Psychol. Monogr., 1941, 23, 191-234.

6. Beeson, M.F. Intelligence at senescence. J. Appl. Psychol., 1928, 4, 219-231.

7. Bellis, C.J. Reaction time and chronological age. Proc. Soc. Exp. Biol. and Med., 1935, 30, 801-803.

8. Binet, A. and Simon, T. The development of intelligence in children. Publications of the Training School at Vineland, New Jersey, 1916, P. 336.

9. Birren, James E. A factorial analysis of the Wechsler Bellevue scale given to an elderly population. J. Consul. Psychol., 1952, 16, 399-405.

10. Brody, M.B. The measurement of dementia. J. Ment. Sci., 1942, 88, 317-327.

11. Brozek, J. Changes in sensory, motor, and intellective functions with age. Geriatrics, 1951, 6, 221-226.

12. Buros, O.K. The Fourth Mental Measurements Yearbook. New Jersey, The Gryphon Press, 1953, P. 338.

13. Chambers, W.G. How words get meaning. Ped. Sem., 1904, 11, 30-50.

14. Chesrow, E.J., Wosika, P.H. and Reinitz, A.H. A psycho-

metric evaluation of aged white males. Geriatrics, 1949, 4, 169-177.

15. Chodorkoff, B. and Mussen, P. Qualitative aspects of the vocabulary responses of normals and schizophrenics. J. Consult. Psychol., 1952, 16, 43-47.

16. Corter, H.M. Factor analysis of some reasoning Psychol. Monogr., 1952, 8.

17. Cuff, N.B. Social status and vocabulary. J. Genet. Psychol., 1935, 46, 226-229.

18. Dorland, W.A. The triumphs of maturity: I. The age of the master work. Welfare Mag., 1927, 18, 1307-1329.

19. _____. The triumphs of maturity: II. The achievements of maturity which the world might have passed. Welfare Mag., 1927, 18, 1444-1466.

20. Elwood, M.I. A preliminary note on the vocabulary test in the revised Stanford-Binet scale, Form L. J. Educ. Psychol., 1939, 30, 632-634.

21. Eysenck, M.D. An exploratory study of mental organization in senility. J. Neurol. Psychiat., 1945, 8, 15-21.

22. Feifel, H. Qualitative differences in the vocabulary responses of normals and abnormals. Genet. Psychol. Monogr., 1949, 39, 151-204.

23. _____. An analysis of the word definition errors of children. J. Psychol., 1952, 33, 65-77.

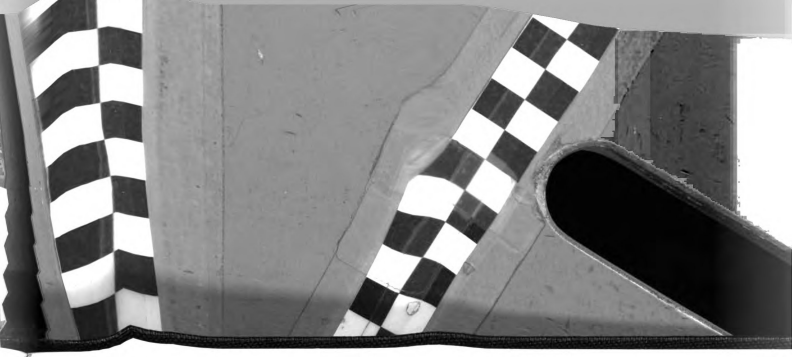
24. _____, and Lorge I. Qualitative differences in the vocabulary responses of normal children. J. Educ. Psychol., 1950, 41, 1-18.

25. Foster, J.C. Significant responses in certain memory tests. J. Appl. Psychol., 1920, 4, 143-154.

26. Foulds, G.A. Variations in the intellectual activities of adults. Amer. J. Psychol., 1949, 62, 238-246.

27. _____ and Raven, J.C. Normal changes in the mental ability of adults as age advances. J. Ment. Sci., 1948, 94, 133-144.

28. Gansl, Irene Vocabulary: Its measurement and growth. Arch. Psychol., 1939, No. 236.



29. Gerstein, Reva A. A suggested method for analyzing and extending the use of Bellevue-Wechsler vocabulary responses. J. Consult. Psychol., 1949, 13, 366-370.

30. Gilbert, J.G. Mental efficiency in senescence. Arch. Psychol., N.Y., 1935, No. 188.

31. _____. Memory loss in senescence. J. Abnorm. Soc. Psychol., 1941, 36, 73-86.

32. _____. Understanding Old Age. Ronald Press, N.Y., 1952.

33. Goldfarb, W. An investigation of reaction time in older adults. New York, Teachers College, Columbia Univ. Contributions to Education, 1941, No. 831.

34. Gray, W. and Holmes E. The development of meaning vocabularies in reading. Publications of the Laboratory Schools of the Univ. of Chicago, 1938, No. 6.

35. Green H.J. A qualitative method for scoring the vocabulary test of the new revision of the Stenford-Binet. Unpublished MA thesis, Stanford Univ., 1931.

36. Greene, E.B. Vocabulary profiles of groups in training. J. Educ. Res., 1940, 33, 569-575.

37. Guilford, J.P. Fundamental Statistics in Psychology and Education. New York, McGraw-Hill Co., 1950.

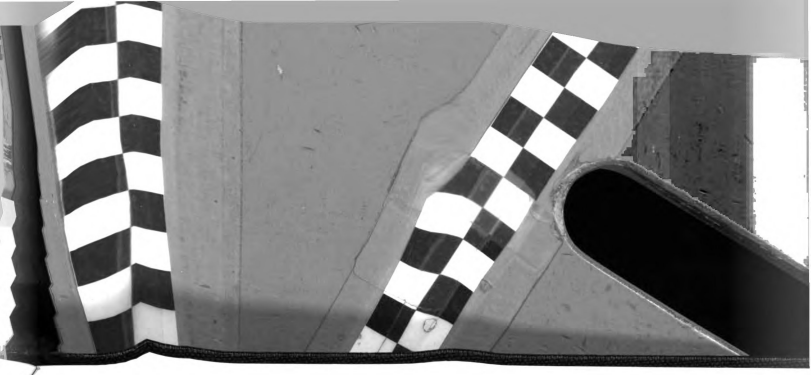
38. Harrington, R. and Ehrmann, J.C. Complexity of response as a factor in the vocabulary performance of schizophrenics. J. Abnorm. Soc. Psychol., 1954, 49, 362-364.

39. Harris, A.J. and Shakow D. The clinical significance of numerical measures of scatter on the Stanford-Binet. Psychol. Bull., 1937, 34, 134-150.

40. Hirsch. E. Uber senile denk und sprachstarungen. Psychol. Forsch., 1928, 10, 358-392.

41. Hunt, W.L. The relative rates of decline of Wechsler-Bellevue "hold" and "don't hold" tests. J. Consult. Psychol., 1949, 13, 440-443.

42. Hunt, J.F. A practical, clinical test for organic brain damage. J. Appl. Psychol., 1944, 28, 175-178.



43. Jones, H.E. and Conrad, H.S. The growth and decline of intelligence: a study of a homogeneous group between the ages of 10 and 60. Genet. Psychol. Monogr., 1933, 13, 223-298.

44. Kaplan, O.J. Mental decline in older morons. Amer. J. Ment. Def., 1943, 47, 277-285.

45. _____. (ed) Mental Disorders in Later Life. Stanford Univ. Press, 1945.

46. Kendig, I. and Richmond. W. Psychological Studies in Dementia Praecox. Ann Arbor Mich., Edwards, 1940.

47. Kirkpatrick, E.A. A vocabulary test. Quoted by W.S. Gray and E. Holmes. The Development of Meaning Vocabularies in Reading. Univ. Chicago Press, 1938.

48. Klopfer, W.G. Personality patterns of old age. Rors. Res. Exch., 1946.

49. Kubo, Y. Mental and physical changes in old age. J. Genet. Psychol., 1938, 53, 101-118.

50. Kuhlen, R.G. Psychological trends and problems in later maturity. In Pennington, L.A. and Berg, I. A. (eds.) An Introduction to Clinical Psychology., N.Y., The Ronald Press Co., 1949.

51. Lawton, G. Mental abilities at senescence: a survey of present day research. J. Appl. Psychol., 1938, 6, 607-619.

52. Lehman, H.C. The creative years in science and literature. Sci. Mon., N.Y., 1936, 43, 151-162.

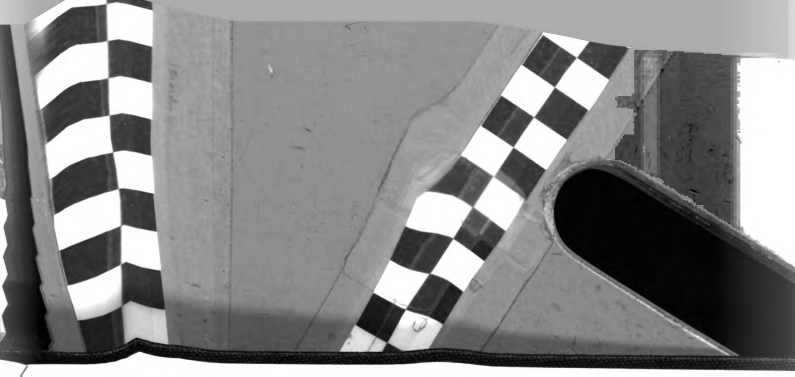
53. _____. The creative years: medicine, surgery and certain related fields. Sci. Mon., N.Y., 1941, 52, 450-561.

54. _____, and Camertsfelder, W.S. Man's creative years in philosophy. Psychol. Rev., 1942, 49, 319-343.

55. Lewinski, R.J. Vocabulary and mental measurement, a quantitative investigation and review of research. J. Genet. Psychol., 1948, 72, 247-281.

56. Lorge, I. Mental status as a function of mental test. Amer. J. Orthopsychiat., 1940, 10, 56-59.

57. Madonick, M.J. and Solomon, N. The Wechsler-Bellevue in individuals past 60. Geriat., 1947, 2, 34-40.



58. Magaret, A and Simpson. M.M. A comparison of two measof deterioration in psychotic patients. J. Consult. Psychol., 1948, 12, 265.

59. Marx, B. A study of the first fifty words of the Stanford-Binet vocabulary. Unpublished MA thesis, Stanford Univ., 1928.

60. Miles, W.R. Age and kinephantom. J. Gen. Psychol., 1934, 10, 204-207.

61. _____. Measures of certain human abilities throughout the life span. Proc. Nat. Acad. Sci., Wash., 1931, 17, 627-633.

62. _____. Abilities of older man. Person. J., 1933, 11, 352-357.

63. _____. The maintenance of our mental abilities. Sci. Mon., N.Y., 1933, 34, 549-552.

64. _____. Age and human ability. Psychol. Rev., 1933, 40, 99-123.

65. Moran, L.J., Moran, Francis. A., and Blake R.R. An investigation of the vocabulary performance of schizophrenics: II. Conceptual level of definitions. J. Genet. Psychol., 1952, 80, 107-132.

66. Prados, N. and Fried, E.G. Personality structure of the older age groups. J. Clin. Psychol., 1945, 3, 113-120.

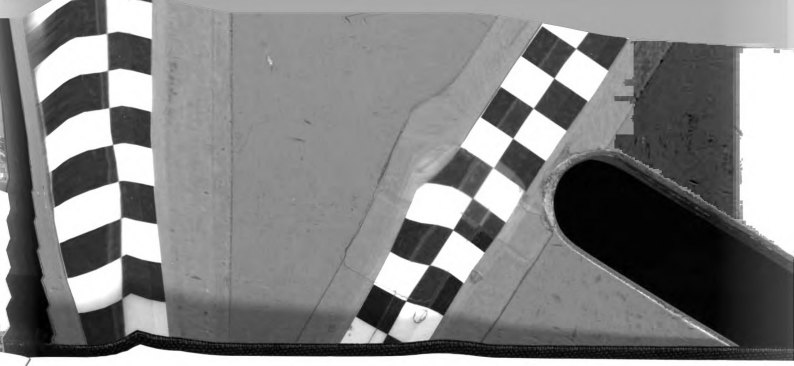
67. Rabin, A.I. Vocabulary and efficiency levels as functions of age in the Babcock method. J. Consult. Psychol., 1947, 11, 207-211.

68. _____. Psychometric trends in senility and psychosis of the senium. J. Gen. Psychol., 1945, 32, 149-162.

69. _____, King, G.F. and Ehrmann, J.C. Vocabulary performance of short-term and long-term schizophrenics. J. Abnorm. Soc. Psychol., 1955, 50, 255-258.

70. Rappeport, D. Diagnostic Psychological Testing. Chicago, Year Book Publishers, 1945, Vol. 1.

71. Reichard, S., Schneider, N. and Rappaport, D. The development of concept formation in children. Amer. J. Orthopsychiat., 1944, 14, 156-161.



72. Rochwarg, H. Changes in the structural aspects of perception in the aged; an analysis by means of the Rorschach test. Unpublished Ph.D. thesis, Michigan State College, 1954.

73. Roe, A. and Shakow, D. Intelligence in mental disorder. Ann. N.Y. Acad. Sci., 1941, 42, 361-490.

74. Rorschach, H. Psychodiagnostics. (3rd ed.). N.Y., Grune on Stratton, 1942.

75. Ruch, F.L. The differentiative effects of age upon learning. J. Gen. Psychol., 1934, 11, 261-286.

76. Schwesinger, G.C. The significance of vocabulary in the interview. Psychol. Clin., 1930, 19, 123-130.

77. Shipley, W.C. A self administering scale for measuring intellectual impairment and deterioration. J. Psychol., 1940, 9, 371-377.

78. Simmons, C. Studies in experimental psychiatry. J. Ment. Sci., 1933, 79, 704-734.

79. Spache, G. The vocabulary tests of the Revised Stanford-Binet as independent measures of intelligence. J. Educ. Res., 1943, 36, 512-516.

80. Stacey, C.L. and Portnoy, B. A study of the differential responses on the vocabulary sub-test of the Wechsler Intelligence Scale for Children. J. Clin. Psychol., 1950, 6, 401-403.

81. Terman, L.M., and Merrill, M. Measuring Intelligence. Boston, Houghton Mifflin Co., 1937.

82. United States Bureau of the Census; Report prepared by the Bureau of Labor Statistics, United States Department of Labor, Washington, D.C.

83. _____; Current population reports, United States Bureau of the Census. 1951, Washington, D.C.

84. _____; Employment problems for older workers. United States Bureau of Labor Statistics, 1950, Washington, D.C.

85. Watts, A.F. The Language and Mental Development of Children. London and Toronto, Harrap., 1944.

86. Weschsler. D. The Measurement of Adult Intelligence. Baltimore, The Williams and Wilkens Co., 1944.

87. Weisenburg, T., Roe, A., and McBride, D.I. Adult Intelligence. N.Y., Commonwealth Fund, 1936.

88. Yacorzynski, G.K. An evaluation of the postulates underlying the Babcock deterioration test. Psychol. Rev., 1941, 48, 261-267.

89. _____. An evaluation of the postulates underlying the Babcock deterioration test. Psychol. Bull., 1940, 37, 425-426.