Reports from the Research Laboratories

of the

Department of Psychiatry

University of Minnesota

The Porteus Mazes: A Critical Evaluation

by

MARY RIDDLE

and

ALAN H. ROBERTS

Report Number PR-74-3

July 15, 1974

The Porteus Mazes: A Critical Evaluation

Mary Riddle and Alan H. Roberts

University of Minnesota

July 15, 1974

PR-74-3

Acknowledgment

This project was supported in part by the Psychiatry Research Unit, Department of Psychiatry, University of Minnesota; Social and Rehabilitation Service Research and Training Grant Number 16-P-56810; and a Grant-in-Aid from the Graduate School, University of Minnesota.

Preface

This is a book about a single psychological test - the Porteus Mazes. It is written by psychologists, but it is intended for all professional people who might be interested in intelligence, impulsiveness, foresight and planning ability, and the effects that these variables might have upon human behavior. In particular, these variables are likely to be related to the ability to succeed in a job, delinquency, criminality, recidivism, the effects of brain damage, and to the evaluation of the social or productive skills of individuals with generally low intelligence.

In evaluating the potential effectiveness of the Porteus Mazes in these areas of human functioning we have read and critically evaluated every study we could locate relating to this test. We report not only what the authors of the various studies tell us but also our own evaluation of the effectiveness and meaningfulness of their work. In many instances data have been reanalyzed and re-evaluated.

This book is written at a time when psychological testing is generally being called into question on both social and moral, as well as methodological grounds, but also at a time when effective objective measurement of personality is needed as an aid to understanding and alleviating many of our social problems. The American Psychological Association is presently rewriting and re-evaluating its Standards for Educational and Psychological Tests and Manuals last published in 1966. Psychologists and others are becoming deeply sensitized to the strengths and the shortcomings of present psychological evaluation methods and their use and misuse.

As the book will show, too much has been claimed for the Porteus Mazes while, at the same time, the test is able to do more than most people give it credit for. On the basis of our critical evaluation we believe that the Porteus Maze Test has great potential as a measurement tool in many of the areas mentioned above. Only a careful and objective examination of the evidence could permit us to be convincing about this.

A very large amount of work by many people has contributed to the preparation of this book. Without Stanley D. Porteus, of course, there would have been no book at all, and during its preparation he provided much in the way of supportive comments and advice. John Brantner first inspired us to look more carefully at the Mazes and he, together with Starke R. Hathaway, Edna Maneval, Manfred Meier, and Lloyd Sines offered many critical comments, suggestions, pats on the backs and nearly slanderous comments all of which made this a better book. Darleen Kenyon and Ella Larson were responsible for typing and proofreading most of the chapters, and Joanne Garcia provided other assistance with respect to the preparation of this volume.

The work involved in the preparation of this book was supported, in part, by Social and Rehabilitation Services Grant RT-2 to the Department of Physical Medicine and Rehabilitation of the University of Minnesota, the Psychiatry Research Special Budget provided by the State of Minnesota, and by a research grant from the Graduate School of the University of Minnesota. To all of these people and organizations we offer our thanks.

Mary Riddle

Alan H. Roberts

University of Minnesota
February 15, 1973

Table of Contents

Chapter I

Introduction

Introduction

There has developed, in the past few years, increasing professional and public concern about psychological assessment. Issues have been raised about the invasion of privacy and about the use of assessment procedures to unfairly discriminate against members of various groups such as minorities, women, or the culturally disadvantaged. Sometimes it is said that the tests themselves are unfair. At other times it is felt that the tests are misused so as to unfairly discriminate.

In the context of these concerns, it is appropriate that a thorough, unbiased, critical evaluation be made of all psychological assessment procedures. But it is particularly appropriate that this be done for the Porteus Mazes since this test would seem to be especially suited for the assessment of many disadvantaged groups. For one thing, it is non-verbal and even the instructions can be given in pantomine. The advantages that one group mav have over another with ABC's may have little relationship to performance on the Maze Test.

Secondly, the Porteus is claimed to test areas of performance and function that are different from those usually measured by intelligence tests. If a person performs poorly on one of the usual measures of "intelligence" because of cultural disadvantage, that person is just as likely to perform poorly in anglo-saxen, middle class schools for the same reasons. One of the major uses of intelligence tests is to predict success in school. A low score on one of the better IQ tests will indeed predict poor school performance even if the person tested is assumed to have some kind of

"innate intelligence" not being properly measured by the instrument.

Our interest in the Porteus Mazes was first aroused when one of the authors served as a consultant to an American Indian resident boarding school in the Southwest. We received a large number of consultation requests to evaluate students who were all described somewhat as follows: "This student is doing very poorly in school but we feel that he ought to be doing better." When we asked the teachers to explain why the student "ought to be doing better", we were invariably told that he was personable, bright, intelligent, alert, had good planning ability, and was otherwise a potentially good student. They couldn't understand why he or she was doing so poorly in class.

When we administered the standard individual intelligence tests to these American Indian students in this school situation which was modeled upon white, middle class methods and values, the students would invariably get scores in the low dull normal or borderline ranges of intelligence. Their scores on the IQ tests were low enough to suggest that they would do exactly what they were referred for doing - performing poorly in school.

However, when the Porteus Mazes were administered to these same students, they invariably scored extremely well on the test. The Maze scores were in the range expected for someone who might be personable, bright, intelligent, alert with good ability to plan ahead, and who would otherwise give the impression of being a potentially good student. The Maze Test was apparently picking up a kind of intelligence or ability which was not helping them

to do well in reading, writing, or arithmetic but which was notable and noticeable to the teachers. The other IQ tests originally constructed to predict school success made them appear to be dull. They clearly possessed talents which were apparently overlooked by the usual IQ measures.

These kinds of clinical observations and impressions surely deserve empirical verification since it is quite probable that "low IQ" is not the only reason children fail to succeed in our schools. There is certainly more than one kind of intelligence, and if the Maze Test or any other test is able to evaluate skills and talents needed for success or even survival in our complex society, these assessment techniques should be evaluated and used widely and wisely.

The Maze Test deserves special attention for yet another reason. It has been claimed for the test that it is a reliable and valid measure of the construct variously described as impulsiveness, ability to delay gratification, future time perspective, etc. In operational terms, this means that the test has potential for predicting behaviors which are of great immediate social concern: delinquency, criminal behavior, recidivism, job stability, and perhaps the potential for drug usage or the capacity to profit from programs designed to help those using drugs.

While it is clear that no single test, or even clusters of tests, can solve the many problems in these areas, any assessment devices with the potential to predict behaviors in these areas should be fully explored and evaluated. As will be shown, many

claims have been made about the Maze Test as a tool for assessment in these areas. If the test does what has been claimed for it, it should be used much more widely than it is now. If on the other hand, the Porteus fails to live up to these claims, the failures should be stated openly and clearly.

Data on the Maze Test have been gathered by researchers and clinicians for over six decades. Sufficient information should be available to assess this assessment device. The potential importance of the technique clearly justifies a thorough, systematic critical evaluation.

John P. Brantner of the University of Minnesota has spoken of the virtues of the Porteus Mazes in his unpublished lectures:

> The Porteus Maze Test comes closer than any other to being the ideal or perfect psychological test. Brief, quickly and easily administered, it samples a "single" behavior enthusiastically performed, leaving a permanent record reliably scored. This record, and, indeed, its two scores make fairly accurate predictions in each of the areas of practical concern to the clinician--intelligence, brain damage and personality. In a minor way, perhaps, and yet very significantly it speaks about the subject's vocational and social adjustment; it describes an important part of his personality and temperament; it reveals his motivational system; it generally displays his approach to problem situations; and it exemplifies a central aspect of his life style. It can be used with the hampered and handicapped; it does not require language; it reaches across cultures; it serves children

and adults. It can only be a combination of the test's deceptive superficial simplicity and professional perverseness that has hindered its widespread study and use.

Many others have agreed with these sentiments in whole or in part (e.g., Burt, 1921; Cunningham, 1916; Fry & Pulleine, 1931; Porteus, 1915, 1924, 1950, 1958; Tizard, 1951). Enthusiastic acceptance of the test by even the most difficult of subjects is indicated in this excerpt from the field notes of Kenneth David made while testing Australian aborigines:

> Testing seemed to create a special bond, and people would wander up and talk and offer boomerangs, and wherever we went ex-subjects would wave and fraternize. I would recommend that any anthropologist who wanted to establish good rapport in minimum time should take a bundle of mazes with him. And I am sure it wasn't just the cigarettes. It became such a common phenomenon that we ceased remarking on it in our field notes, but we would invariably give an S a cigarette before testing started. After a few puffs he would put his cigarette down in the sand and forget about it, and it would simply burn itself into a long stick of ash. Ss sat through the dinner bell, the end of working-day whistle, to finish a testing trial, or ignored the invitations of their friends to go elsewhere or of their superiors to go to work. During one memorable testing, a fight developed about a hundred yards away, and the whole camp congregated there, to adjudicate between the contestants, or help out, so that it was a near riot. This attracted the Superintendent, who came roaring over

> in his car; but our two Ss merely looked up, shook their heads, and went on with their work [Porteus, Bochner, Russell, & David, 1967, p. 10].

If the test indeed possesses the characteristics claimed, it should be one of the most popular, most frequently used of the available psychometric devices. However, in the two most recent surveys on the frequency of use for various psychometric instruments in several different types of psychological services (Lubin, et al., 1971; Sundberg, 1961), the Maze test was ranked only in the middle range. Sundberg (1961) found the Maze test to be most frequently used in Veteran's Administration hospitals, where it ranked 17th out of 31 tests, and least frequently used in University-affiliated clinics, where its frequency of administration was among the lowest of the tests reported in the survey.

Either the test's qualities are exaggerated, or the Mazes deserve a better reception than they have had among psychologists. Critical evaluation would appear worthwhile. However, the only systematic attempt to review the Maze test literature, other than the reviews published from time to time by the test's author, was done by Tizard (1951), 20 years ago. That review was not particularly detailed or critical, and a great many studies have been added to the Maze test literature in the past 20 years. The purpose of this book is to fill this hiatus by reviewing the rather massive literature, critically evaluating it, and where meaningful, reanalyzing and reinterpreting the data.

The Maze test was created by Porteus in 1913, and was first published in 1915 in the Journal of Experimental Pedagogy and the Journal of Psycho-asthenics. As Porteus is quick to point

out, the test has long outlived both of these journals. Since their first publication, the test instructions have gone through a number of revisions (Porteus, 1919, 1924, 1933, 1942, 1950, 1955, 1959a, 1959b, 1965).

In addition, a number of other investigators have published their own instructions for the Maze test (Arthur, 1930, 1943, 1947; Bronner, et al., 1927; Burt, 1921, 1948; Gaw, 1925a; Herd, 1930). The net effect of the abundance of different sets of Maze test instructions seems to have been to encourage investigators to deviate from standard administration procedures.

There are currently three separate series of Porteus Mazes, the original or Vineland Revision Mazes, and two new series which have been developed by Porteus to eliminate the effects of practice for the purpose of retesting: The Extension Series (Porteus, 1955) and the Supplement Series (Porteus, 1959b).

On the basis of the subject's performance on any of the series of mazes, two scores are derived: A Test Age (TA) or Test Quotient (TQ), and a Qualitative Score (Q). Maze TA, as it is currently scored, is based on the level, in years, assigned to the highest maze passed in the allowed number of trials, minus one-half year for each unsuccessful trial occurring at that level or below. After TA has been calculated, it is sometimes converted into TQ by means of a table provided by Porteus. The relationship between TA and TQ on the Maze test is analogus to the relationship between MA and IQ on the various forms of the Binet Intelligence Test. Both TA and TQ are often referred to as "quantitative score" in Maze test literature. In this paper we use "TA" to designate the quantitative score unless there is a good reason to

refer to TQ specifically. The Q-score is based on the weighted sum of the number of certain technical errors occurring during maze execution. The errors that are scored include lifting a pencil contrary to instructions, crossing lines or cutting corners, changing direction, and other stylistic "errors" that are not directly related to the correct execution of a maze. Confusion often results because TA and Q-score are scored in opposite directions. It should be emphasized that a high TA indicates superior performance while a high Q-score indicates inferior performance.

In addition to TA and Q-score, a number of other scoring systems have been used with the Maze test. Investigators have reported the speed with which subjects trace through the Mazes (Foulds, 1951, 1952; Foulds & Caine, 1958, 1959a, 1959b; Petrie, 1952; Porteus & Bassett, 1920; Shapiro, et al., 1958; Shapiro, et al., 1960 Shapiro, et al., 1962; Tow, 1955; Vernon, 1937; Verrill, 1959; Waite & Neilson, 1919) or the length of a subject's pause between the time he is given a maze blank and the time he begins to trace through it (Foulds, 1951, 1952; Foulds & Owen, 1964; Shapiro, et al., 1958; Shapiro, et al., 1960; Shapiro, et al., 1962; Sheer, et al., 1952).

Foulds devised a method of administering the Mazes under "distraction". Distraction consists of the examiner counting while the subject traces through the mazes repeating each number after the examiner. Foulds felt that this method of test administration, in combination with results from the Maze test administered in standard fashion, might be useful in psychiatric diagnosis. However, it appears that subjects, regardless of their

diagnosis, perform faster and begin maze execution sooner under distraction conditions than under standard conditions (Foulds, 1952; Foulds & Caine, 1958, 1959a, 1959b; Shapiro, et al., 1958; Shapiro, et al., 1960; Shapiro, et al., 1962).

Gruetzow and Brozek (1946) and O'Shea, et al., 1942) used a system of scoring Maze performance based on the number and difficulty of the mazes successfully executed. Foulds' Weighted Wrong Direction (WWD) score is based on the number and length of blind alley entrances (Foulds, 1951). Porteus (1959a, 1961, 1962) developed a Conformity-Flexibility (C-F) score (later [Porteus, 1955] renamed Conformity-Variability [C-V]) based on certain similarities between two different tracings of the year XI maze by the same subject. However, because so little work has been done using these various atypical scoring systems this critical review will deal almost exclusively with TA and Q-score.

A number of investigators (e.g., Burt, 1921; Porteus, 1918, 1919; Porteus & Babcock, 1920; Tizard, 1951) have mentioned that observation of a subject's reactions during his execution of the Maze test gives insight into certain aspects of his personality. There is no published research of this claim.

There are a number of articles in the literature dealing with mazes other than the Porteus Mazes (Ball, 1929; Beach & Johnston, 1960; Blackwood, 1927; Covault, 1955; Davies, A. & Davies, M. G., 1965; Davies, M. G., & Davies, A., 1965; DeSanictis, 1931; Elithorn, 1955; Elithorn, Jones, Kerr & Lee, 1964; Elithorn, Kerr& Jones, 1963; Elithorn, Kerr & Mott, 1960; Gibson, 1955; Husband, 1930; Kunce & Worley, 1966; Matarazzo, et al., 1955; Perkins, 1927; Peters, 1956; Pomeroy, 1950; Rafi, 1960;

Scott, 1936; Smith, 1927; Tachibana- 1931). None of the mazes described in these articles have been extensively investigated, and no attempt has been specifically made to relate findings with these mazes to the Porteus Mazes. However, it may be noted that none of the published data for these mazes in any way contradict Porteus Maze data. On the contrary, findings for mazes even remotely resembling the Porteus Mazes are generally supportive of Porteus Maze findings.

The history of the Maze tests' development is briefly reviewed, followed by the main body of this book, which is organized in terms of the basic considerations that face the psychologist when he uses a psychological test. These include the characteristics of the test itself such as norms, reliability and administration; the influence on test scores of the subject characteristics that are already known to him such as race, sex, age, social class and whether medication is being administered; and the interpretation of the test findings in the light of what is known about the test's validity. This paper deals with each of these three basic considerations separately. In most areas investigations of both TA and Q-score are available, and in these cases the TA discussed is first followed by a discussion of Q-score findings. There are a few areas of study that are exclusively the domain of TA (planning capacity and social adjustment) or of Q-score (anxiety and stress).

Chapter II

History of the Maze Test and Development of Maze Test Instructions

History of the Maze Test and Development of Maze Test Instructions

The Maze test had its beginning in 1913 when Porteus was responsible for selecting students to be placed in a special training school for the feebleminded in Melbourne, Australia. After observing the performance of children labeled feebleminded on various tasks, Porteus observed that foresight was an important variable for which no test had been devised. The Maze test was developed after other means of measuring foresight and planning ability had been tried and rejected.

The Porteus Mazes were first published in 1915 in the Journal of Experimental Pedagogy and the Journal of Psycho-Asthenics. At the time of its first publication, the test was entitled the "Motor-Intellectual Series." This name, however, did not survive.

The Vineland Revision, developed while Porteus was working in the training school at Vineland, New Jersey, was first published in 1919, although this series of mazes did not acquire its present form until 1933. The changes in Maze test administration, scoring and form prior to 1933 were made largely to improve standardization. Since 1933, the changes in Maze test application were mainly due to modifications that permitted additional uses for the test.

The first version of the Maze test consisted of eleven consecutive mazes starting with year III and ending with year XIII. The instructions for administration and scoring in the original articles were quite brief and lacked specificity in comparison to later versions of the test. Only a general description of the administration was given, but no really specific instructions were included. The final score for the original version of the test

was a test age score; this was simply the age corresponding to the highest maze passed. All the mazes in the series were to be given regardles of failure. Some of the mazes in this original series are identical to those currently in use; others are only slightly different in form, and a few of the original mazes were subsequently eliminated from the series. Over all, the test was probably easier in 1915 than it presently is, largely because the scoring was more lenient and more demonstration was given by the examiner. The maximum possible score, however, was 13 years, indicating that for older subjects TAs would be lower, on the average, than they would be as the test is administered today.

In the early Maze test literature, probably as a result of the lack of specificity of the instructions, the methods of administration described by different authors were quite varied. It can be assumed that some of these differences must have affected the test results. One common variation, which undoubtedly affected the results, was the practice of using a single maze blank for each year, rather than presenting the subject with a fresh maze as soon as an error was made. The subject was allowed to continue tracing through the maze until it was completed, and each blind alley entrance occurring on the maze was scored as an error. The calculation of the final test age score was based on the number of errors made on each maze rather than the number of trials used in each year. This practice would tend to make the test easier and would be expected to result in higher TAs than would be obtained using the current method of administration.

In 1919 Porteus published a new set of Maze test instructions in his booklet Porteus Tests-The Vineland Revision. Between 1915

and 1919, the test underwent changes in format. Some of the mazes were changed to different age levels, some were eliminated, and some altogether new mazes were added.

From Porteus' later description (e.g., Porteus, 1965) of the changes occurring in the mazes between 1915 and 1919, this appears to have been quite an active period of transition. Many changes were tried, but most of these were unpublished and could not have been used in studies employing the mazes. Most of the changes incorporated into Porteus Tests-The Vineland Revision were the result of findings with samples of normal school children. The mazes for the lower years appeared to be somewhat too easy for these children. Young children tended to obtain Maze TAs higher than their chronological ages. The original year XIII maze appeared to be too difficult and became the year XIV maze. Porteus discovered that standardization was improved by not including a year XIII maze in the series.

The instructions in Porteus Tests-The Vineland Revision were more explicit. This probably made the administration of the test by different examiners more uniform. Requirements for passing some of the mazes became stricter, as did the scoring of the final test record with the introduction of deductions for Maze test errors and the cessation of testing after failure in two successive years. These changes would be expected to lower the final TAs; however, the maximum possible TA was raised one year, which should have increased the final score. Because these two factors appear to have opposite effects, it is difficult to predict their combined influence on the final TA. For older subjects, because of the 14-year ceiling, and for very young, low

scoring subjects, TAs would probably be lower, on the average, than would be expected today, while for young subjects TAs would be higher. Some of the mazes were a little simpler in design than they are today.

In 1924, Porteus brought out new Maze test instructions in Guide to the Porteus Maze Test. Two changes in the general format of the test were described in this publication. These were the addition of two adult tests and the description of retesting by using inverted mazes. As in 1919, the instructions for Maze test administration became more explicit and complete. The range of possible scores became greater with this version of the test.

The exact wording to be used during the administration of each maze was made explicit in the instructions in the Guide to the Porteus Maze Test. The specific wording to be used in administration had only been given for the year V maze in the 1919 Vineland Revision instructions. The range of possible test scores was broadened by stating that any attempt, however crude, to follow the outline of the year III maze indicated ability at the two year level, and by the fact that the maximum TA obtainable became 18 years, when the adult mazes were given. No restrictions were put on their use.

The two adult mazes were added to the series to make the test applicable at higher age levels. Inverted test instructions were included for the readministration of the entire series so that retesting could be possible. Porteus first wrote about this use of the test in 1923, but no specific instructions for application were provided at that time. Porteus states that early in the development of the test he believed that the method of

administration could be rather flexible and that differences in administration techniques would not significantly affect the resultant scores. Further investigation of the Maze test indicated the need for a more uniform, standard method of application (Porteus, 1950). This led to the specifications of the exact wording to be used in test administration.

The scoring changes introduced in Guide to the Porteus Maze Test would be expected to raise TA, on the average. The instructions were made more explicit, which should have resulted in more uniform test administration. Because the test had a higher ceiling and several of the mazes were simpler than is currently the case, TAs for subjects of all ages, except the very young, would be expected to be somewhat higher using the maze test as applied in 1924 than they would be using the 1965 version of the test. Because of the stricter scoring of the mazes for years III and IV in 1924, very young, low scoring subjects would probably obtain lower scores than would be expected now.

In 1933 a new set of Maze test instructions was brought out by Porteus in The Maze Test and Mental Differences. In this publication, the Vineland Revision mazes were revised for the last time. The maximum TA score obtainable on the test became 17 years, with the elimination of the second adult maze. Tables for determining a test quotient (TQ) based on Maze TA were first presented in this publication. TQ was designed to give a score that would be comparable to an IQ score from an intelligence test. An explanation of the derivation of the TQ table was also presented. The table represents a modified form of the usual age-score ratio, mental age divided by chronological age multiplied by 100. The

table is modified in that the divisor used in the calculation of TQ was not always the exact chronological age.

Porteus said that 20 years of experience with the Maze test led to most of the changes in the maze patterns and in the scoring presented in The Maze Test and Mental Differences. Prior to their publication, the revised mazes were administered to one thousand Honolulu school children. This work was done by Dr. Russell Leiter, who incidentally, was responsible for adding the lines to the year V Maze. Apparently, Porteus was impatient to get his manuscript for the 1933 revision on an early boat. In those days only two boats a week went to the mainland. He made a last minute decision to add the lines, and at the time Leiter was the only one available to do the work. The lines were hastily applied using a ball point pen and metal ruler, which, Leiter explains, is why they are so uneven.

Porteus (1933) wrote that the changes made improved standardization, although he did not present the data on which this statement was based. Inspection of the changes suggests that their overall effect was to make the test more difficult. Also, the maximum possible TA was one year lower than it had formerly been, so TAs would, in general, be lowered in comparison to their expected level in 1924.

For subjects below 13 years of age the 1933 version was probably somewhat more difficult than the current version, due to the stricter rules governing termination of the test in 1933. It is difficult to predict the relationship between the 1933 and 1965 versions of the test with respect to adults. Stricter rules for test termination applied to adults in 1933, but adults were given

the added practice of working through the mazes for years V and VI in 1933, while adult testing now begins with the year VII maze, which should have tended to make the 1933 test easier. Thus, TAs based on the 1933 version of the test could be expected to be roughly equivalent to TAs based on the current versions.

In 1942, Porteus published Qualitative Performance in the Maze Test in which Q-score was presented for the first time. Some minor changes in administration were necessitated by the introduction of Q-score, although most of the rules given in The Maze Test and Mental Differences were still followed. In 1933, the warnings to be given to the subject concerning pencil lifts were quite strict. These instructions had to be changed with the introduction of Q-score. Most of the changes in administration were for the year V maze. The first and only warning about pencil lifts was to be given with this maze. It was emphasized that the examiner should draw his demonstration line very carefully, keeping it in the middle of the path and making a ninety degree turn at the first corner. A warning not to cross lines or enter blind alleys was added to the instructions to be given to the subject during the administration of the year VII maze. The rules for administration of the remaining mazes were not revised in 1942.

The weights to be assigned for each Q-score error were given in the instructions along with a description of each type of error. The Q-score was to be calculated from the performance on the mazes for years V-XIV only. The addition of a Q-score in no way changed the scoring of TA. The rules for determining TA described in The Maze Test and Mental Differences still applied in 1942, although a TA penalty for pencil lifts included in the

1933 scoring was dropped. On the sample mazes for years V-XIV, provided with the instructions, an example of a wavy line was drawn for the purpose of comparison, and arrows marking the first and last thirds of each maze were added.

Differences in the quality of Maze test performance among different types of subjects had long been observed by Porteus and others. Porteus thought that these differences were due to temperamental traits, and that a consistent, quantified measure of some of these differences would be useful. Fourteen elements of poor qualitative performance, including those which finally became the components of Q-score, were counted each time they occurred in the maze records of a group of delinquents and a group of nondelinquents. Those styles of performance differentiating the delinquents from the nondelinquents were included in the Q-score. Those which did not discriminate were eliminated. On the basis of their ability to discriminate the maze performance of delinquents from that of nondelinquents, weights were assigned to the components of Q-score. It appears from Porteus' descriptions that these weights were assigned on an intuitive basis from inspection of the relative frequencies of the Q-score errors occurring in the two groups. The ability of the weighted Q-score to discriminate criminals and delinquents from normals was then checked by comparing the Maze test records of 100 delinquent boys and 100 delinquent girls to the records of 100 nondelinquent boys and 100 nondelinquent girls as well as comparing the Maze test records of 100 criminals to those of 100 bus drivers. The qualitative scores of the delinquents and criminals were over twice as high as those of the normals.

It does not appear that the changes in administration included in the Qualitative Performance in the Maze Test would have materially affected the final test age score. It is possible that the elimination of the test age penalty for pencil lifts would tend to raise TA slightly, on the average, but in general TAs obtained from this version of the test would be expected to be equivalent to TAs from the 1933 version. The overall effect of the differences between Q-score in 1942 and the present was probably to make Q-scores somewhat lower using the 1942 instructions, at least among relatively low scorers. A combination of the following factors made the Q-score lower in 1942; demonstration by the examiner, a warning against crossing lines on the year VII maze, the clear specification that no more than one wavy line error was to be scored for each year, the fact that the instructions implied that Q-score was only to be based on those mazes used in the scoring of TA, not including mazes above the ceiling, and the fact that the adult maze was not included in the scoring. The inclusion of the Q-scores obtained on the years V and VI mazes in 1942 might tend to increase Q-score, but not enough to overcome the effect of the other variables. The setting of an absolute upper limit on Q-score in the 1965 instructions would tend to make the 1942 scores higher for very high scoring subjects than they are currently.

Very few changes were incorporated into the instructions in The Porteus Maze Test and Intelligence brought out in 1950. This set of instructions essentially combined the quantitative portion of the test, as described in 1933, with the qualitative portion of the test as described in 1942. Also, an alternative set of

instructions was provided in The Porteus Maze Test and Intelligence. These instructions were written by Dr. Henry N. Peters and could be used in place of the instructions given by Porteus for testing adults.

There was no change in the scoring of TA introduced in The Porteus Maze Test and Intelligence. The descriptions of the various types of Q-score errors were briefer in 1950 than they were in 1942, and it was no longer clear that only one wavy line error could be scored at each year. Some changes were made in the scoring of pencil lifts. After 10 pencil lifts, the subject was to be warned again to keep his pencil on the maze. After this second warning, all pencil lifts were to be penalized five points each. The maximum weighted score to be recorded for pencil lifts was 80. Previously no limit was set on this score. The examiner was to make a note if the Q-score was determined from less than nine mazes.

Combining the instructions from the 1933 and 1942 versions of the test probably clarified the proper method of administration. The level of TA would still be expected to be equivalent to its level using the 1933 version of the test. While the change in lift pencil scoring and the ambiguity of wavy line scoring present in 1950 would tend to elevate Q-scores somewhat above their level using the 1942 instructions, Q-scores were probably still slightly lower on the average than they are at present.

Shortly after The Porteus Maze Test and Intelligence came out, Porteus published The Porteus Maze Test Manual (1952) in England. The Maze test instructions in the latter publication were an almost verbatim duplication of the instructions in the

former, differing only in the use of one inconsequential word.

In 1955 the Extension Series of mazes was introduced in The Maze Test: Recent Advances, which presented a complete list of instructions for the application and scoring of the Vineland Revision, as the original series was called, and Extension Series. The instructions published in The Maze Test: Recent Advances gave a brief list of the changes to be made from the instructions in The Porteus Maze Test and Intelligence, which was still to be the major source of test instructions. The Extension Series was intended only for subjects aged 14 years and over. In 1955, the testing was to start with the year VII maze for older subjects in both the Vineland and the Extension Series. The mazes for years V and VI could still be used for subjects below age 14 or for demonstration purposes with subjects having a language handicap. The pencil lifts warning was to be given after five such errors on a single maze as well as after a total of 10 pencil lifts in the series. Prior to the 1955 instructions warning was only given after 10 pencil lifts.

The entire series of Vineland Revision mazes was to be given to its completion, but this change did not affect the TA scoring. TA was still calculated on the series of mazes completed prior to two successive failures or any three failures. Successes in subsequent mazes were not considered in TA calculations.

A limit of 48 weighted points was set as the maximum possible score in each of the Q-score categories as well as a limit of 100 total points for Q-score. The Q-score was to be based on all the mazes administered, including the Adult maze. This was the first time the Adult maze had been used in the calculation of the Q-score.

Q-score errors occurring on the year VII maze were to be penalized an extra point. Formerly this was done on the years V and VI mazes.

The Extension Series was created to be used in retesting, and was intended to follow the administration of the Vineland form of the mazes. Designed to eliminate the practice effects of having previously taken the other form of the Maze test, the Extension Series was, essentially, a mirror image of the Vineland mazes with a few changes in structure to make each maze somewhat more complex. This series of mazes was purposely made more difficult than the original series in order to compensate for practice so that the two series of mazes, when administered in the proper sequence, yielded comparable scores. The changes in the structure of the mazes were made on an intuitive basis. The appropriateness of these changes was confirmed by studies in which the two series were consecutively applied, and they were found to yield comparable scores (Porteus, 1955).

The subject was to be allowed to complete the entire series of mazes to make the Q-scores on different administrations more equivalent. The Q-score was to be based on the performance on all the mazes. The reason for warning the subject about pencil lifts after he had committed this error five times on one maze was that some subjects begin a maze by sketching rather than by drawing a continuous line. Porteus believed these subjects should not have to be penalized thirty points for failing to understand the nature of the task they were to perform. The examiner was to interrupt this type of maze execution sooner than was previously allowed.

The starting maze for the administration of the Vineland Series of mazes was changed to year VII for older subjects because this was the lowest year of the Extension Series and starting at the same level would make the administrations of the two series more similar.

The changes in Maze test administration introduced in 1955 made the test very similar to the form in current use. However, only three trials were allowed on the adult mazes in 1955, and the rules for test termination were somewhat stricter in 1955 than they are at present. These factors have tended to lower TA slightly, on the average. Basing Q-score on the entire series of mazes as was done in 1955, regardless of the point at which the TA ceiling was reached, would tend to elevate Q-score above the level expected today since the number of mazes on which Q-score is to be based is currently rather arbitrary. The added warning after five pencil lifts on any maze would be expected to decrease Q-score, but probably not sufficiently to compensate for the effect of basing Q-score on all the mazes. Overall, Q-score was probably somewhat higher in 1955 than it is at present.

The Maze Test and Clinical Psychology (1959) was mainly a combination of the instructions in The Porteus Maze Test and Intelligence and the changes suggested in The Maze Test: Recent Advances. This 1959 publication served to make the 1955 instructions more explicit and also included a few minor changes of its own. For either series, the set of mazes to be used for adults, years VII through Adult I, was to be given to all subjects age 13 years and older. Formerly the lower age limit was 14 years.

It was no longer required that the entire series be administered, regardles of failure. The examiner was not to draw a line around the first turn in the year V maze. Instead, he was only to indicate the starting place by pointing with his finger. This same change was made in the instructions for the year VII maze to be used with older subjects. The weighting of pencil lifts was no longer to increase after 10 lifts. The weight was to be three points for each pencil lift regardless of frequency. Q-score errors on the maze for year VI were to be penalized an extra point for children. Formerly this penalty was given for errors on mazes for year V and year VI. It was no longer specified that the Q-score was to be calculated from the performance on the entire series of mazes years VII through Adult I. This method of Q-score calculation was required in 1955, but was made optional in 1959. A new table for converting test age scores to TQ scores was given in The Maze Test and Clinical Psychology. A sample for scoring wavy lines on the Adult maze was provided.

Many of the differences between the instructions given in The Porteus Maze Test and Intelligence and those given in The Maze Test and Clinical Psychology were due to the addition of the Extension Series and most of these changes were suggested in The Maze Test: Recent Advances. Another source of change was the idea that the style of maze execution may reveal aspects of the subjects personality. Several changes were made so that the presentation of the mazes would be as unstructured as possible to avoid interfering with the subject's own style of maze execution. This accounts for most of the changes in administration for the year V and year VII mazes.

The main difference in test administration between The Maze Test and Clinical Psychology and the current version of the test was the rather unstructured presentation of the mazes to the subject to avoid interfering with his own style of maze execution. This change may have tended to increase the Q-score. Previously, Porteus (1942) stressed the importance of demonstrating careful maze execution so that the subject would follow the examiner's example and do a careful job of tracing through the mazes, which would improve his Q-score.

A new TQ table was introduced in this publication. This is the table currently in use. In general, the effect of this table was to tighten the distribution of TQ scores. Subjects with high TQs within each age group would have lower TQs on the new table while subjects with low or medium TQs would have higher TQs on the new table. Only those subjects with moderately high TQs had scores that were relatively unchanged on the new table.

The Supplement Series of mazes was introduced in 1959 in a little booklet entitled The Porteus Mazes: The Supplement Series of Mazes. No changes whatever in test administration were made at this time.

The most recent set of Porteus Maze instructions are included in Porteus Maze Tests: Fifty Years Application published in 1965. In these instructions it states that timid and withdrawn subjects were allowed to have practice trials on the mazes for years V and VI. Primitive subjects and subjects with language problems were also allowed practice trials on these two mazes. The practice trials were to be continued until the subject was able to execute the year VI maze correctly without any assistance from the examiner. When testing a primitive subject or a frightened child, it

was stated that a familiar person could be present in the testing room during the demonstration and practice portion of the test. If this person stayed for the administration of mazes year VII and above, he was to be seated so as not to be able to see the subject's performance, and he was not to be allowed to speak.

More complete instructions for the testing of primitive subjects were provided in Porteus Maze Tests: Fifty Years Application than in previous publications.

The 1965 instructions state that testing was to begin with the year III maze for subjects under six years old. Previously the year III maze was the starting point for children under five. The examiner was to demonstrate the execution of the maze by drawing a line just around the first turn and then give the subject a new year III maze to do. Previously, the examiner was to draw about one inch along the path and then let the subject continue on the same maze. Three trials, rather than two, were to be allowed on this maze. Three trials, rather than two, were to be allowed on the year IV maze as well. After the second trial, the examiner could point out all the places where a line was crossed or touched. This maze was considered to be passed if no more than three lines were crossed. Formerly the requirements for passing this maze was no more than two crossed lines. The examiner was to draw approximately one and one-half inches along the path before giving the year V maze to the subject. In The Maze Test and Clinical Psychology, the examiner was not to draw on the year V maze during the administration of the test. He was only to indicate the starting point with his finger.

The instructions for the year VI maze no longer stated that

the examiner was to indicate the starting point with his finger. Instead he was instructed to refer to the rat and the cheese in his presentation of this maze to the subject.

The instructions for the year VII maze also no longer stated that the examiner was to indicate the starting point with his finger. He was permitted to indicate the exit of the year VII maze to very primitive subjects. The administration of the remaining mazes was essentially unchanged.

The instructions for scoring TA were given in greater detail. The scoring in the 1965 instructions system was not changed from the previous instructions. The wording of the Q-score instructions implied that each Q-score error was to be penalized an extra point on both the year VI and the year VII mazes for children. Formerly this was done only on the year VI maze.

Many of the changes put forth in The Maze Test and Clinical Psychology, to assure that the examiner would not unnecessarily influence the subject's style of maze execution, were undone in Porteus Maze Tests: Fifty Years Application. The former instructions allowed almost no form of demonstration on the part of the examiner. The 1965 instructions were more moderate with respect to demonstration and quite similar to the Maze test instructions prior to 1959. The more complete instructions for testing primitive subjects came from Dr. James Gregor, who described the techniques he used in the testing of Australian aborigines. The leniency in TA scoring introduced for the III and IV year mazes in 1965 would tend to elevate the scores of very young subjects somewhat over the level to be expected at any time since 1915, when the scoring was even more lenient than it is now.

The previous instructions for administering the Maze test, from 1915 through 1959, were examined and impressions were formed regarding their comparibility to the current instructions. These impressions are based on speculation about the effect on TA of the various differences between the current Maze test administration instructions and those of the past. Unfortunately, data are not available to test the validity of most of these statements.

1. Using the 1915 instructions for test administration, young subjects would probably be expected to obtain higher scores and adults lower scores than would be the case using the current instructions.
2. Using the 1919 instructions, very young low scoring children would be expected to obtain lower scores; older, brighter children somewhat higher scores; and adults lower scores than they would using the current instructions.
3. The 1924 instructions would be expected to produce somewhat higher scores for all subjects than do the current instructions except for very young low scoring children, who would be expected to score lower with the 1924 rules.
4. The 1933 instructions might produce comparable scores among adults to what would be expected currently, but only if those variables which appear to influence the score are fairly well counterbalanced. For children under 13, scores would be expected to be lower using the 1933 instructions.
5. The 1942 instructions differ from the current instructions in the same manner as the 1933 instructions, and the same relationships among scores would be expected.

6. The 1950 instructions are also comparable to the 1933 instructions.
7. The 1952 instructions are identical to the 1950 instructions.
8. The 1955 instructions would be expected to produce lower scores for all subjects than would the present instructions.
9. Except for very young, low scoring children, who would score lower with the 1959 instructions, subjects would be expected to score the same with the 1959 instructions as with the current instructions.
10. The magnitude of the above differences in scores is unknown, but would probably not be very large in any case.

Similar impressions were formed about Q-score. As with the statements made about TA, the statements about Q-score made here are of a speculative nature. Unfortunately, most of the data that would be required to verify these statements are unavailable.

1. On the average, the Q-scores expected with the 1942 version of the test would be lower than with the current version, except for very high scorers who could obtain higher scores with the 1942 instructions than would be currently possible.
2. Q-scores obtained with the 1950 test instructions would probably be generally higher than those obtained with the 1942 version of the test, but lower than those obtained with the current version.

3. Q-scores obtained with the 1955 rules would tend to be higher than would currently be expected.
4. Q-scores obtained with the 1959 rules would also tend to be higher.
5. The magnitude of the above differences is unknown, but would probably not be very large in any case.

Chapter III

Characteristics of the Test

Characteristics of the Test

Norms

Test Age

With the exception of cross-cultural studies, most of the work on the Porteus Maze Test Age standardization was done prior to 1920. During this period, over 2000 children between the ages of 5 and 14 years were tested and the scores for children at each age were reported (Porteus, 1919). Since these early studies were done, the test has undergone considerable revision. The current Vineland Revision Mazes did not reach their present form until 1933 and the instructions for Maze test administration have been revised several times since then (Porteus, 1942, 1950, 1955, 1959a, 1965). It is unlikely that the results of these early normative studies would apply to the tests as they are presently administered.

In 1921, norms for British children were published by Burt. In Burt's study, the administration of the Maze test was different in some respects from that recommended by Porteus in any of his publications. Burt failed to mention how many subjects were tested in obtaining his normative data. It is difficult to ascertain whether these results would apply to the test in its current form.

The three normative studies, done with children, that come the closest to being of value today are those by Porteus and Babcock (1926), Arthur (1933a), and Havighurst and Janke (1944). The study by Porteus and Babcock and the study by Arthur were done using the 1924 Maze instructions (very slightly modified in Arthur's study), which have greater similarity to the present

instructions than do those used in previous studies. Havighurst and Janke did not identify the source of their instructions, but there is reason to believe they used Arthur's version. The results of these three most recent studies are summarized in Table 1. These results appear to indicate that the test is somewhat too

Insert Table 1 about here

easy, but not grossly so, when administered according to the 1924 instructions. Since the current instructions presumably make the test harder in this age group, the standardization is more likely to be correct, but these are the most recent studies attempting to standardize the Mazes. The test, as currently administered, has never been standardized.

A number of studies have been done that can be used to arrive at adult norms. Seven studies have reported Maze test results for subjects age 14 to 65 who do not appear overly deviant (Arthur, 1933a; Porteus, 1955; Porteus & Babcock, 1926; Roberts & Alcuaz, unpublished; Small, 1954; Sterne, 1969; and Weisenberg, Roe & McBride, 1936). These studies involve 16 groups and a total of 652 subjects. The mean TAs reported for these groups range from 13 to 15 years, the majority being within 1/2 year of 14. The five groups of subjects with a mean score of 15 years were all under 30 years of age while the two groups with a mean score of 13 years tended to consist of older subjects. One of the low scoring groups was made up of subjects 50 to 59 years old and the other of subjects with a mean age of 41. On the basis of these studies, it would seem reasonable to assume that adult subjects

Table 1

TA Standardization

Investigator	Porteus & Babcock (1926)			Arthur (1933)						Havighurst & Janke (1944)			Total	
Subjects	American School Boys			American School Children						All 10-year-olds in a Midwest Comm.				
				Male			Female							
Age	N	M	SD	N	M	SD	N	M	SD	N	M	SD	N	M
5:0- 5:11	---	---	--	56	6.38	1.57	40	5.88	1.35	---	---	--	96	6.16
6:0- 6:11	---	---	--	62	7.34	1.34	53	6.79	1.86	---	---	--	115	7.09
7:0- 7:11	---	---	--	50	8.28	2.02	52	7.63	1.66	---	---	--	102	7.95
8:0- 8:11	---	---	--	44	9.66	1.99	60	8.63	2.52	---	---	--	104	9.07
9:0- 9:11	30	9.4	--	50	11.32	2.48	50	11.52	2.59	---	---	--	130	10.95
10:0-10:11	30	10.93	--	56	12.10	2.66	44	11.93	2.55	110	12.5	--	240	11.69
11:0-11:11	31	11.38	--	43	13.02	2.44	61	12.81	2.75	---	---	--	135	12.55
12:0-12:11	32	12.16	--	49	13.22	2.63	56	12.64	2.62	---	---	--	137	12.74
13:0-13:11	31	13.03	--	49	13.50	2.06	48	13.43	2.56	---	---	--	128	13.36

can be expected, on the average, to obtain TAs of 14 years, although there may be some tendency for younger subjects to score slightly higher and older subjects slightly lower. Stated differently, any subject scoring 14 years on the Porteus Maze test cannot be said to be abnormal on the basis of this score.

The standard deviations of Porteus TA have been reported for 10 of these same groups. They range from 1.4 to 2.9 years, indicating that the standard deviation of TA in a normal adult group approximates two years. Unfortunately, these results are byproducts of studies that were not really intended to investigate TA norms. No rigorous study of adult TA norms has ever been undertaken.

Arthur (1947, p. 20) points out that the recent popularity of mazes for entertainment has probably influenced the Maze test norms.

> The scores for the Revised Porteus Maze Test were seriously affected by the use of mazes in the comic journals as amusement devices. This is not a mere assumption; it is based on the repeated comment, "These are fun! I do them in the funnies." Ratings calculated with maze and without maze indicate that this influence begins to be significant at the nine-year level.

The possibility that increased experience with mazes in the general population has affected TA norms suggests a need for new norms.

Test Age Conclusions

1. As currently standardized, the Maze test may yield TAs that are close to correct, but research, particularly recent research, is so inadequate that no definite conclusions on this point are possible.

2. The average TA in the general adult population is probably quite close to 14, although research is inadequate to establish this with any certainty.
3. The standard deviation of TA is probably close to two years in the normal adult population.
4. The lack of normative data on a test of such apparent potential utility as the Porteus Mazes is appalling.

Q-Score

Q-score results for 17 groups of normal subjects are reported in the literature (Docter & Winder, 1954; Fooks & Thomas, 1957; Gibbens, 1958; Porteus, 1942, 1945, 1955, 1961; Sanderson, 1945). The weighted mean of these group Q-scores is 22 for a total of 1307 subjects. The group Q-scores range from a mean of 13.66 for a group of Hawaiian high school students (Porteus, 1945) to 25 for Docter and Winder's (1954) ninth and tenth grade boys. By and large, subjects in these studies were adolescents, although a few adult results are reported and these appear to be in agreement with the adolescent findings.

Additional normative data have been published in the form of percentile distributions (Jensen, 1961; Porteus, 1950). Jensen's subjects are quite atypical for inclusion in a discussion of norms, being a group of Air Force inductees who failed initial group testing, but his results show such close agreement with Porteus' results that they are included here for comparative purposes. The subjects in the studies by Porteus are described only as nondelinquents.

Table 2

Q-score Percentiles

Percentiles	Jensen's Airmen N=634	Porteus' 1942 Subjects N=826	Porteus' 1948 Subjects N=1200
0	186	90	102
10	54	46	47
20	39	35	36
30	30	28	29
40	24	24	24
50	20	20	19
60	17	16	15
70	13	13	12
80	9	9	10
90	7	6	7
100	0	1	0

Insert Table 2 about here

Q-Score Conclusions

1. The average Q-score in the normal population appears to be approximately 22, but research was based mainly on the performance of adolescents.
2. Additional information for adults is needed.
3. The variability of Q-score among normals appears to be quite large.

Reliability

Test Age

There are 11 studies from which various types of reliability estimates for TA can be obtained (Arnold, 1951; Dentler & Mackler, 1962; Docter, 1960b; Ebert & Simmons, 1943; Gibbens, 1958; Morgenthau, 1922; Porteus, 1918, 1924, 1955; Weintraub, 1968; Wittenborn & Mettler, 1951). One interscore reliability, of 1.0, has been reported for TQ (Weintraub, 1968). It is not surprising that two trained examiners were in perfect agreement on TQ, since no subjective judgment is involved in TQ scoring.

There is one test-retest, one alternate form and two Kuder-Richardson reliabilities in the literature, that used young children as subjects, and that appear to be uncorrected for age. These correlations range from .75 to .97 with a mean (using the Z-score conversion method of calculation) of .89 for a total of 126 subjects. Because there is a strong relationship between age and TA among children, the correlations may be unrealistically high.

There are three test-retest correlations available in which age is controlled: .43 for 97 subjects (Gibbens, 1958), .66 for 47 subjects (Porteus, 1918) and .75 for 20 subjects (Wittenborn & Mettler, 1951), and there are two alternate form reliabilities available; both are .50 (Docter, 1960b; Porteus, 1955) for 60 and 50 subjects respectively. The results indicate that .5 is about the magnitude of the reliabilities to be expected using either of these methods of estimation.

Unfortunately, the test-retest and alternate form methods may not fairly assess the actual reliability of the Maze test. Porteus has frequently theorized that individuals differ in their ability to profit from experience, and, since this ability is supposed to be measured by the Maze test, high test-retest reliability would indicate low validity.

The studies by Ebert and Simmons (1943), Morgenthau (1922) and Porteus (1924) were not used in the above calculations. In the former study, the method of test administration almost surely invalidated the results. In the latter two studies, it was unclear whether age was controlled.

Test Age Conclusions

1. Failure to control for age appears to inflate estimates of Maze TA reliability.
2. Test-retest and alternate form reliabilities approximate .5. These methods of reliability assessment may be inappropriate, however, for the Maze test.

Q-score

Nine studies report interscorer reliabilities for Q-score (Anthony, 1963a, 1963b; Crown, 1952; Docter & Winder, 1954; Fooks & Thomas, 1957; LaBarba, 1965; Ruebush, et al., 1963; Sarason, et al., 1960; Schalling & Rosén, 1968; Weintraub, 1968). From their results, it appears that interscorer agreement is quite good. The reliabilities reported range from .89 through .99.

The alternate form reliabilities reported in the literature are not high, .75 (N=96) in Briggs' (1963) study and .51 (N=60)

in Docter's (1960b) study. A test-retest reliability of .59 (N=97) was reported in Gibbens' (1958) study. These reliabilities suggest that Q-score should only be used to make rough classifications and not fine discriminations among subjects.

Q-score Conclusions

1. Interscorer agreement, with respect to Q-score, is quite high.
2. Q-score is probably not sufficiently reliable to make fine discriminations among subjects.

Comparability of Different Forms

Test Age

There are five studies in which TAs for both the Vineland Series and the Extension Series are reported for the same subjects (Dentler & Mackler, 1962; Docter, 1960b; Porteus, 1955; Porteus & Barclay, 1957; Porteus & Gregor, 1963). A total of 19 groups involving 519 subjects were used in these studies. The greatest difference between means within any group for the two series of Maze tests was 1.08 years for a group of 29 retarded children (Dentler & Mackler, 1962). Porteus & Gregor (1963) reported a difference of one year for a group of 12 adult female Australian aborigines. For all other groups studied the differences between means were less than a year. None of the differences between the Vineland means and the Extension means were significant. Subjects represent a variety of groups: High school students, adult and adolescent Australian aborigines, retarded children, psychotic psychiatric patients, and schizophrenic Army enlisted men. Group membership did not appear to

influence the difference between Vineland and Extension means.

In two studies (Porteus, 1955; Porteus & Gregor, 1963) both series of mazes were administered on the same day; in another (Docter, 1960b) there was an interval of three days between testings. The test-retest interval was reported as one month or less in one of the studies (Dentler & Mackler, 1962) and as six weeks or more in the remaining study (Porteus & Barclay, 1957). The length of the interval between testings did not appear to affect the difference between means on the two series of mazes, at least not for the relatively short intervals investigated. The implication of these studies is that within any group the Vineland and Extension Series will give equivalent scores, when properly applied, even for intertest intervals as long as six weeks.

The Supplement Series has not been as widely investigated as the Extension Series. Only two studies compare scores on the Supplement Series with scores on the other two series of mazes (Porteus, 1959b; Porteus & Gregor, 1963). The subjects in Porteus' 1959 study were a group of high school students, 249 males and 259 females. Only the Extension and Supplement Series were administered. The intertest interval was not reported, but both series were undoubtedly administered in the same testing session for each group. There was less than 0.1 years difference between the Extension TA and the Supplement TA. Porteus and Gregor (1963) reported mean scores on all three series of mazes for four groups of Australian aborigines. They also reported mean scores on the Extension series and the Supplement Series for a group of high school boys who admitted to some history of trouble with the law. In all cases the mean TA for the Supplement Series differed by less than a

year from the mean TA for either of the other series of mazes. None of the differences were statistically significant. All three series of mazes were applied on the same day to the aborigines. There was no mention of the intertest interval of the high school boys.

Test Age Conclusions

1. When subjects are tested with the Extension Series within approximately six weeks following the administration of the Vineland Revision, the mean TAs on the two series of mazes can reasonably be expected to be equivalent. The effect of longer intertest intervals has not been investigated.
2. The Supplement Series tests probably give TA scores that are equivalent to those on the other two series of mazes. Additional research investigating the similarity of scores on the Supplement Series to scores on the other series is needed.

Q-score

Four studies compare Q-score performance of subjects on the Vineland Revision mazes with their performance on the Extension Series (Docter, 1960b; Fooks & Thomas, 1957; Porteus, 1955; Porteus & Gregor, 1963). While there were no significant differences in Q-score between the two series of mazes reported for any of the 11 groups of subjects investigated in these studies, certain meaningful changes have been reported in two of the studies comparing an initial administration of the Vineland mazes with the Extension Series. Docter(1960b), in an investigation of the test-

retest performance, with a three-day interval, of 60 hospitalized schizophrenics, reported a mean Q-score of 32.8 on the Vineland Revision which dropped to 23.7 on the Extension Series. This difference failed to reach statistical significance but it is important to note that the subjects' scores changed from the "delinquent" level to the "nondelinquent" level. Fifty-five percent of the subjects obtained Q-scores above 29 on the Vineland Revision while only 25% had scores in the "delinquent" range on the Extension Series.

Fooks and Thomas (1957) tested a group of 50 delinquents and 50 nondelinquents using the Vineland and Extension mazes. Again, there was no significant change in Q-score for either group on the second application of the test. Both versions of the test correctly identified well over 70% of the delinquents. On the Vineland Revision, 28% of the nondelinquents were misidentified as delinquent by the test, which is typical of the false positive rate found in other studies. On the Extension Series, however, 42% of the nondelinquents were misidentified. This percentage increased to 54% using unweighted Q-score and a cutoff point of 16. In this study, the Extension Series Q-score was of questionable validity as a predictor of delinquency.

Among the eight groups totaling 349 subjects investigated in the two studies by Porteus (Porteus, 1955; Porteus and Gregor, 1963), the Vineland and Extension Series Q-scores were closely matched, and no findings were reported that would suggest any differences between the two series with respect to Q-score.

Only two studies have been done reporting Q-score retest data using the Supplement Series (Porteus, 1959b; Porteus & Gregor, 1963). Both studies report mean Supplement Series Q-scores that

do not differ significantly from Q-scores on the Extension Series. In the study by Porteus and Gregor (1963), all three series were given in succession. Subjects were 49 adult male Australian aborigines. None of the differences between Q-scores on the three series was significant. The 508 high school students in the 1959 study by Porteus were not tested with the Vineland Revision. No study of the equivalence of the three series, administered as recommended by Porteus, using subjects from any Western culture has ever been published.

Q-score Conclusions

1. No significant differences in Q-score among the three series of Porteus Mazes have been reported in the literature.
2. Certain unexplained changes, from the Vineland to the Extension mazes, in the proportion of subjects found within the delinquent range of Q-score require additional investigation.

The Effect of Changes in Administration

Test Age

A few studies provide data on the effects of altering some aspect of test procedure on the resultant TA (Arthur, 1930, 1933, 1943, 1947; Briggs, 1963; David, 1967; Ebert & Simmons, 1943; Gibbens, 1958; Porteus & Gregor, 1963; Shakow & Millard, 1935).

Both Arthur (1930, 1947) and Ebert and Simmons (1943) report normative studies involving children, ages 5 through 15 years, in which the Maze test was administered providing added practice. Using Arthur's (1930) Form I norms as a basis for comparison

(this version of the mazes closely approximates Porteus' 1924 version), the norms for the Maze test with added practice were at least half a year higher than the Arthur Form I norms in 19 out of a total of 27 possible comparisons. Compared to the composite of all the normative studies done within this age range in which the test was administered in standard fashion (see Table 1), the norms for the Maze test involving added practice were at least half a year higher in 21 out of the 27 possible comparisons.

Arthur (1933b) reported lower mean TAs at each year of chronological age for subjects age 5 through 15 years when mazes were inverted after a previous failure, as recommended by Porteus, than were obtained by the subjects when this procedure was not followed. Briggs (1963) compared the TA for 96 subjects using their dominant hand with their TA using the nondominant hand and found no significant difference between hands. Practice was controlled in this study through the use of a counterbalanced design. David (1967) found no significant difference in TA between a group of 23 Australian aborigine children who were verbally reinforced for correct Maze performance and 23 aborigine children who were not reinforced.

Four studies contain information on the effect of including additional mazes at the upper end of the test (Arthur, 1933b; Gibbens, 1958; Porteus & Gregor, 1963; Shakow & Millard, 1935). Arthur (1933b) gave test results from subjects age 5 through 15 years, with the two adult mazes excluded from the scoring. The exclusion of these two mazes did not affect the mean scores for subjects age 5 through 7 years, but at every age level for

subjects 8 through 15 years, the mean TA was lower without the additional mazes. The difference between mean TAs with and without the adult mazes increased with the age of the subjects from 0.29 years of TA at age 8 to 1.02 years of TA at age 15.

When the Maze TQ in the study by Gibbens (1958) was based on all the mazes through Adult I regardless of earlier failure, rather than terminating testing according to Porteus' instructions, TQ increased from 94.0 to 99.2. Including the Adult I Maze in the scoring produced TA increases of 0.6, 1.4 and 0.3 for three groups of aborigines as compared to TA calculated without the Adult I Maze (Porteus & Gregor, 1963). When the Maze test results for 81 adult subjects (Shakow & Millard, 1935) were scored according to Porteus' 1924 rules the mean TA was 14.3. The TA dropped to 13.2 when the 1933 scoring was used. This probably resulted because the 1924 version included an additional adult maze and had a ceiling one year higher than the 1933 version.

These results all accord with expectation. Many other modifications of Maze administration have been tried (e.g., Arnold, 1951; Frost, 1965; Herd, 1923, 1930; Porteus, 1955, 1959a), but there is no information currently available to assess the influence of these modifications on TA. Interpretation of Maze test results is most valid when the test has been administered in the standard fashion. It is unfortunate that standard administration has not been routinely followed, since a wealth of information has been lost due to unnecessary modifications.

Test Age Conclusions

1. Giving added practice when administering, the Maze test tends to increase scores.
2. Inverting a maze after a previous failure, as recommended by Porteus, decreases TA.
3. Use of the nondominant hand in testing should not materially affect the TA results, on the average.
4. Including additional mazes at the top of the test tends to increase TA.

Q-score

The variables related to the effect of Maze administration procedure on Q-scores that have been systematically investigated are the use of the nondominant versus dominant hand to execute the Mazes (Briggs, 1963), the number of mazes administered (Docter & Winder, 1954; Gibbens, 1958), the relationship between individual Q-score errors and total Q-score (Barry, et al, 1961; Docter, 1960a) and the weighted system of scoring (Docter & Winder, 1954; Gibbens, 1958; Grajales, 1945; Porteus, 1942; Sanderson, 1945; Schalling & Rosén, 1968; Wright, 1944). Briggs (1963) found that significantly higher Q-scores were obtained by subjects using their nondominant hand in comparison to their dominant hand scores. Docter and Winder (1954) found a significant correlation of .29 between Q-score and number of mazes required by their 60 delinquent and 60 nondelinquent subjects to complete the entire series of mazes. Gibbens (1958) reported that when Q-score was based on the entire series of mazes, his subjects scored 35.1 and when it was based on only those mazes used to calculate maze TA, the mean Q-score was 33.5.

Some investigators have gone to elaborate pains to modify Q-score so that it would be based on the same number of mazes for all subjects (e.g., Anthony, 1963a, 1963b; Leibowitz, 1966; Peters & Jones, 1951). These modifications hardly seem justified on the basis of the small effect the number of mazes administered actually seems to have on Q-score. These idiosyncratic scoring systems only serve to make comparisons between studies more difficult, and do not particularly enhance the value of the individual study.

Two studies investigated the relationship between total Q-score and individual Q-score errors (Barry, et al., 1961; Docter, 1960a). In both studies, the number of pencil lifts gave the highest correlation with total Q-score, .83 (N=70) and .63 (N=224) respectively. Barry, et al., found that the sum of pencil lifts and crossed lines correlated .9 with total Q-score for each of four separate groups of subjects. The correlations of other errors with total Q-score reported in these studies were all too low to be of any practical utility, although many of them were significant.

A number of investigators (Docter & Winder, 1954; Gibbens, 1958; Grajales, 1945; Porteus, 1942; Sanderson, 1945; Schalling & Rosén, 1968; Wright, 1944) reported both weighted and unweighted Q-scores for their subjects. The data suggest that weighted Q-score can be estimated with a reasonable degree of accuracy by doubling unweighted Q-score, for group data. This finding is discussed in greater detail in the section on delinquency, criminality, and psychopathy.

Q-score Conclusions

1. Use of the nondominant hand, as opposed to the dominant hand, tends to increase Q-score, indicating a lack of equivalence between Q-scores based on the nondominant and dominant hand.
2. The evidence that is available concerning the effect of the number of mazes administered on Q-score indicates that it is not large, accounting for approximately 9% of the Q-score variance. This area needs further investigation to determine whether the consternation it apparently causes some investigators is justified.
3. Correlations of lift pencil and lift pencil plus crossed line with total Q-score are sufficiently high so that results reported for these Q-score errors may be considered suggestive of total Q-score trends in the reported direction.
4. Weighted Q-score can apparently be estimated by doubling unweighted Q-score, for group data.

Practice

Test Age

A number of studies report TAs (or the difference between TAs) for initial testing and retesting of subjects with an intertest interval of less than one year (King, 1949; Landis & Erlick, 1950; Peters & Jones, 1951; Porteus & Barclay, 1957; Porteus & Peters, 1947; Sheer, 1956; Sheer, et al., 1952; Wittenborn & Mettler, 1951). These studies involved a total of 213 subjects. For the most part these subjects were psychiatric patients. All

investigators reported an increase in score on the second testing except Sheer, et al., who found a decrease of .25 years in one group and Wittenborn and Mettler who found a decrease of 1.89 years. The weighted mean increase in scores from the first to the second testing for all 23 groups of subjects reported in the literature is 1.25 years of TA.

There appears to be a relationship between the increase in TA on retesting and the initial level of TA. Twelve groups with a total N of 77 had an initial mean TA below 12 (Jones & Peters, 1952; Landis & Erlick, 1950; Sheer, 1956; Sheer, et al., 1952), while 9 groups with a total N of 59 had initial mean TA of 12 or more (King, 1949; Landis & Erlick, 1950; Sheer, 1956; Sheer, et al., 1952). The groups with initial TAs below 12 showed an average increase of 2.0 years upon retest while groups with initial TAs of 12 or more increased only 1.1 years upon testing. On the basis of these data, it appears that the lower the score on the first application of the Mazes the greater the increase in score on the second application. Unfortunately, it is not possible to construct an adequate test of the significance of this trend. It would be more appropriate to divide subjects on an individual basis into groups obtaining high and low scores on the first application of the Mazes. None of the studies reported in the literature published individual scores so this type of analysis is impossible at present.

It is also of interest to determine whether the same relationship between TA and practice gains is found when performance on the second application of the Maze test is compared to performance on the third. Two groups with a total of 42 subjects

had a mean TA below 12 on the second administration of the Maze test (Landis & Erlick, 1950; Porteus & Barclay, 1957) while 7 groups involving 57 subjects had mean scores above 12 years (King, 1949; Landis & Erlick, 1950; Sheer, 1956; Sheer, et al., 1952). The groups testing below 12 on the second administration showed an average increase in TA of 1.0 year from the second to the third administration of the test while the groups scoring above 12 showed an average increase of 0.5 years on the third administration. The mean gain in TA for all groups was 0.7. The improvement from the second to the third application of the test does not appear to be as pronounced as the improvement from the first to the second, however, the general pattern is similar. It appears that the effects of practice are greatest when the TA scores are lowest.

The only study in which performance on the third application of the Maze test was compared to performance on the fourth application was done by Sheer, et al., in 1952. The group of six control subjects in this study had a mean TA of 13.3 on the third application of the test and gained one year of TA on the fourth application. No comparable information on lower scoring subjects is available. King (1949) also reported results for a fourth testing and Landis, et al., (1949) gave the results of a fifth testing, but the previous test results of these subjects were not given.

Other studies are available regarding practice effects on the Maze test. Ruebush, et al., (1963) report practice effects for two groups of fifth grade school boys. There are no Maze scores included in the article, and the actual amount of increase in score on the second testing is not reported.

Doll, et al., (1932) report marked gains in Porteus TA on retesting nine retarded, birth-injured subjects. Unfortunately, for some of the younger subjects the practice effect is compounded with increases in TA due to age since the retesting was done one-half to two and one-half years after initial testing. For the five adult subjects in this study the mean TA on initial testing was 8.3 years (SD=2.2). At the second testing the mean TA was 12.4 years (SD=2.9) for a difference of 4.1 years. Four of these subjects were retested one to two years after initial testing. The fifth subject was retested 7.4 years after initial testing. In spite of this long delay he gained 6 years in TA. Doll's findings are consistent with other findings in the area of practice effects. The unusually large practice gains are paired with an unusually low initial TA. The type of subjects used in this study, however, may not be comparable to other subjects used in the investigation of practice effects and five subjects is an exceptionally small number.

Smith (1960) investigated the effects of practice over an extended period of time using subjects originally tested by Sheer (1956). He administered the Maze test to two groups of subjects for the fourth time 8 years after the third administration of the test. One group of 12 subjects lost 1.8 years of TA from the third to the fourth testing. The other group of 11 subjects lost 2.88 years. These losses brought each group back to just below their level at the first administration of the Maze test. On the basis of this study it appears that the effects of practice are not necessarily permanent. Porteus and Kepner (1944), however, reported a mean gain of 1.74 years of TA for a group of subjects

retested after an average of 4.4 years (SD=3.7).

The apparent inconsistencies in these studies of long term practice effect may be due to aging. In Smith's study, the group showing the greatest loss had a mean age of 59, and the other group had a mean age of 42. Doll's subjects were all younger than 40 years of age throughout his study. The majority of Porteus and Kepner's subjects were adolescents. There were two subjects in this study who were over 40 and both showed losses in TA. It is possible that older subjects show losses in TA with increased age and these losses offset practice gains. With the exceptions of Doll (1932), Porteus and Kepner (1944), and Smith (1960), no investigator reported an interval between tests longer than four months, although some investigators did not specify the length of the interval. Inspection of the available data does not suggest any relationship between intertest interval and amount of improvement with practice for intervals of four months or less.

Test Age Conclusions

1. On the average, gains of about 1.25 years may be expected on the second application of the Maze test due to the effect of practice, at least among psychiatric patients.
2. On the third administration of the test, practice gains are not as large, 0.7 years, and it is probably reasonable to assume that additional testing will produce increasingly less practice effect, in the type of subjects studied.
3. It appears that on any given administration of the test, the lower the initial TA, the greater the expected gains due to practice.

4. The practice effect does not appear to be influenced by the length of relatively short intertest intervals.
5. The effect of long intertest intervals on practice gains is unclear. Some of the reported results may have been complicated by TA losses due to aging.
6. It should be emphasized that generalization of these conclusions to subjects other than psychiatric patients would be risky.

Q-Score

Very little work has been done with the effects of practice on Q-score. By and large, it is assumed not to exist and is ignored. From the few relevant studies (Jones & Peters, 1952; Landis & Erlick, 1950; Ruebush, et al., 1963; Sheer, et al., 1952), it appears that Q-score tends to decrease with practice. Landis and Erlick are the only investigators who actually published mean Q-scores for subjects at each testing. Their findings reflect a group trend toward a decrease in Q-score with each repetition of the Maze test. Subjects in this study were psychiatric patients. Because there are no published data available for any other group, it is impossible to know how accurately these subjects represent psychiatric patients in general or how closely these figures would agree with results from the general population. It would be absurd, on the basis of this one study, to attempt to quantify the magnitude of change in Q-score with each repeated application of the Mazes.

Jones and Peters (1952) reported a decrease in the number of Q-score errors per maze for a group of 11 psychiatric patients

from 5.58 on the first application of the maze test to 4.75 on the second application four months later. It is not possible to calculate the total Q-scores on the basis of the information provided in this article.

Ruebush, et al., (1963) merely mentioned that their 32 subjects, high defensive and low defensive children, showed an overall decrease in Q-score on the second administration of the maze test. However, no scores were reported.

Sheer, et al., (1952) did not analyze total Q-score but did report the mean lift pencil scores per maze for five groups of psychiatric patients. Four of their five small groups showed a decrease in number of pencil lifts on the second administration of the Maze test. One group of two subjects made an increased number of pencil lifts on the second administration. Both Barry, et al., (1961) and Docter (1960a) have published results indication that lift pencil is highly correlated with total Q-score; therefore, this measure can be expected to reflect general Q-score trends.

Out of a total of nine groups studied by all investigators, eight groups showed a change in score on the second administration of the Maze test suggestive of a decreased Q-score.

Q-Score Conclusion

Evidence indicates that the possibility of a practice effect on Q-score must be considered in the evaluation of any study using repeated applications of the Maze test. It is most unfortunate that this aspect of the Maze test has not been more extensively studied.

The Relationship Between Q-score and Test Age

Correlations between Q-score and TA have been reported for nine groups involving a total of 881 subjects (Docter, 1960a; Fooks & Thomas, 1957; Grajales, 1945; Porteus, 1942; Roberts & Alcuaz, unpublished data; Roberts & Erikson, 1968; Wright, 1944). The correlations range from .28 for a group of delinquents in Wright's study, the only positive correlation published, to -.40 for a group of delinquents in Porteus' 1942 study. The mean of the reported correlations is -.22 (using the Z score conversion method of calculation).

Conclusion

There is a low order negative correlation between Q-score and TA.

Chapter IV

The Influence of Subject Characteristics

The Influence of Subject Characteristics

Cross-Cultural Studies

Test Age

The literature concerned with cross-cultural investigations of TA can be divided into studies of nonprimitive and primitive subjects. The various ethnic groups in Hawaii have been the most extensively studied of the nonprimitive groups (Louttit, 1931; Porteus, 1933, 1937, 1939a, 1950, 1961; Porteus & Babcock, 1926). The only consistent finding in these studies has been that subjects of Portugese ancestry tend to get lower scores than other subjects. Porteus found them to have a mean TQ of 86.2 in 1937, 86.65 in 1950, and 91.5 in his 1926 study with Babcock. Using only groups with both sexes combined and arbitrarily defining average performance as TQ=96-104, three out of the four Japanese groups studied gave average performances and one group was above average. Three of four Chinese groups studied scored below average while one group scored above average. Of nine groups of part-Hawaiians in the literature, four scored below average, three were average and two were above average. These studies do not indicate that the ethnic groups in Hawaii differ from white norms. The subjects in these studies were all randomly selected grade school and high school children. Results were also reported by Porteus (1933, 1939a) for children referred to the psychological clinic at the University of Hawaii. These findings are consistent with those reported for school children. None of the studies reported results for adults.

Three articles report investigations of the Maze test performance of American Indian children (Havighurst & Hilkevitch, 1944;

Sparling, 1941; Telford, 1932). In each of these studies the TAs of the Indians were no lower than white norms. Havighurst and Hilkevitch (1944) exercised great care to obtain large and representative samples. They reported the mean TA of subjects at each year of chronological age from 6 to 15 years. Their results are in very close agreement with similar data gathered on white subjects. Scores reported by Sparling (1941) were somewhat higher than white norms, but her small sample of subjects was selected exclusively from children attending school, who, therefore, were probably of somewhat higher intelligence than average. Telford's (1932) sample of 12 twelve-year-olds had a mean TA of 12 on the Maze test.

Zeckel and Van Der Kolk (1939) reported Maze TQs for four groups of normal Dutch children involving a total of 100 subjects. Two of these groups obtained below average scores, one obtained an average score and the remaining group scored above average.

The mean TQ of a subgroup of 30 Spanish-American delinquents did not differ from the mean TQ of the entire sample of delinquents in a study by Roberts and Erikson (1968).

Five studies have been done using American Negroes as subjects (Cooper, et al., 1967; Grajales, 1945; Peters & Jones, 1951; Peterson & Telford, 1930; Weisenburg & McBride, 1935). The studies by Peterson and Telford (1930) and by Weisenburg and McBride (1935) were the only ones using subjects that approximate normals. These two studies report opposite results. Peterson and Telford tested tested 90 twelve-year-old school children living on an isolated island off the coast of South Carolina. These children scored approximately four years of TA below white norms. The

authors speculate that the inferior showing of their subjects on a number of intellectual tests may be due to genetic drift with the brighter inhabitants of the island having gradually migrated to the mainland. Weisenburg and McBride (1935), in a study of hospital patients who were free from disorders known to affect mental functioning, found that their 15 Negro subjects obtained inferior scores on all the intellectual tests administered except the Porteus, on which they were superior to the 70 white subjects used in the study.

The remaining three studies used institutionalized subjects. Grajales (1945), found the 25 Negro and 55 white delinquents he tested were at essentially the same level on the Maze test with mean TQs of 89 (SD=18.92) and 91.36 (SD=17.71) respectively. The two groups of adult hospitalized schizophrenic Negroes studied by Peters and Jones (1951) each obtained TAs of slightly over 10 years, which is well below the white norm, but not at all unusual for groups of schizophrenics. Cooper, et al., (1967) attempted to differentiate Negro institutionalized retardates, who, on the basis of behavior, appeared to be of normal intelligence from those who appeared to be of subnormal intelligence, and found the Porteus to discriminate between these two groups perfectly. This is a strong indication that the test is valid with Negro subjects. The behaviorally normal subjects (N=29) obtained unusually high scores on the Maze test with a mean TQ of 121.7.

Four studies have been done with non-white school children in their native countries (Tsao, 1937; A Report of Intelligence

Testing by the Porteus Maze Test, 1935; MacCrone, 1928; Fick, 1939; Porteus, 1937). The first two studies were done in China using Chinese school children as subjects. One study found the Chinese children to obtain TA scores above the white norms; the other study found them to score below the white norms. MacCrone tested African children, who scored below the white norms. Fick (1939) gave a series of tests, including the Mazes, to African children and concluded that, in general, these subjects test four to five years below white norms. The Maze results were not reported separately, however. Porteus' Formosan schoolboys had a mean TQ of 101.5, but the Formosan schoolgirls only scored 81.3.

Porteus (1950) published the results of testing adult Chinese males in Peiping. Forty literate poorhouse inmates scored 12.0 years, 50 illiterate coolies scored 11.7, and 27 illiterate poorhouse residents scored 10.5. These scores are below white norms, but no comparable whites have ever been studied. Maze results for four other groups of Asiatics were reported by Porteus (1937). Fifty-one illiterate adult male Ainus from Japan had a mean TA of 13 years and 16 female Ainus scored 12.3 years. Two groups of illiterate males from India scored 13.2 and 11.3 years. Some of these scores are quite close to white norms, but again, it is unclear how well whites totally deprived of education would score.

The results using primitive subjects will not be discussed in detail. In general, primitive adults do not come up to white norms. Negritos and East Indian aboriginals appear to make the poorest showing with TAs for adult male subjects ranging from 7 to 9.5 years in the 13 groups investigated (Porteus, 1937;

Ray, 1953; Vicary, 1938). Only one group of adult female Negritos was investigated. These subjects had a mean TA of 6-1/2 (Porteus, 1937). The other primitive adults tested with the Porteus include Australian aborigines, African Bantu, and non-Negrito East Asiatics and Pacific Islanders (DuBois, 1944; Fry & Pulleine, 1931; Gregor & McPherson, 1963; Piddington & Piddington, 1932; Porteus, 1931, 1937; Porteus, Bochner, et al., 1967; Porteus & David, 1966; Porteus & Gregor, 1963). In general, these subjects obtain maze TAs of 10 to 12 for males and 7 to 9 for females. There are a few exceptions to this rule that can be almost entirely accounted for by the method of test administration or subject selection employed in the studies.

The test results of 10 groups of Australian aborigine children have been published (David, 1967; David & Bochner, 1967a; Piddington & Piddington, 1932; Porteus, 1931, 1937; Porteus & David, 1966; Porteus & Gregor, 1963). Five of the 10 groups had average scores on the mazes; the other 5 were below average. It is interesting to note that 4 of the 5 groups scoring below average were composed of part-white subjects.

The question of the effect of education and acculturation on the Maze test results of primitive subjects is complex and cannot be answered on the basis of the information currently available. The authors usually did not report background information necessary for an analysis in this area.

Test Age Conclusions

1. Maze test results have been reported for Portugese, Japanese, Chinese, and part-Hawaiian school children

from Hawaii, as well as for American Indian, Dutch, American Negro, Chinese, African, Formosan and Australian aborigine children. Of these groups, only the Portugese children from Hawaii appear to be below white norms in Porteus performance. Viewed together, the results for the African and American Negro subjects suggest that Negro children may be below white norms in Porteus performance, but further research is needed to establish this finding.

2. Research with nonprimitive adults has been inadequate to determine whether or not racial group differences exist.
3. Primitive adults in general test below white norms with Negritos and East Indian aboriginals obtaining the lowest scores. Sex differences in TA appear to be quite marked among primitive adults.

Q-Score

There is no real evidence that ethnic group membership affects Porteus Maze Q-score, but the studies done in this area are few and their results generally unclear. Two studies give Q-score results for Negro subjects (Grajales, 1945; Peters & Jones, 1951). The subjects in both of these studies were highly selected. Grajales tested institutionalized delinquents, and Peters and Jones tested hospitalized schizophrenics, whose results are further obscured because their Q-scores were not reported in standard fashion. The 25 Negro subjects in the study by Grajales (1945) obtained a mean Q-score of 49.12 points

(SD=19.32), which was 11.75 points lower than the mean of the 55 white subjects. This difference may reflect ethnic differences in the variables affecting institutionalization rather than ethnic differences in Q-score. There was no white control group in the study by Peters and Jones (1951). Their 20 Negro subjects had a mean Q-score of 4.7 per maze, but the number of mazes administered to each subject is unknown, making total Q-score undeterminable. On the basis of the available studies, it is impossible to determine whether the Q-scores of Negroes differ from those of whites.

Roberts and Erikson (1968) reported a study in which the Q-scores of a subgroup of 30 Spanish-American institutionalized delinquents did not differ from the scores of the total group of delinquents made up of the 20 Spanish-Americans, 10 Anglo-Americans, 5 Navajo Indians, and 5 Negroes. While they cannot be assumed to apply to Spanish-Americans as a whole, these results certainly do not constitute evidence for national group differences.

Porteus has published two studies on Q-score within the various national groups represented in Hawaii (Porteus, 1945, 1961). In his 1961 study, he found significant Q-score differences between Chinese, Japanese, and part-Hawaiian tenth grade students. The Chinese scored significantly lower (mean Q-score=20.07) and the part-Hawaiians scored significantly higher (mean Q-score=30.77) than the Japanese (mean Q-score= 26.45). In his 1945 study, the Q-score reported for a group of part-Hawaiian high school students was unusually low, 13.66 for males and 15.82 for females, much lower than the scores reported

for any of his 1961 subjects, while his Japanese and Chinese subjects scored 21.4 and 22.3 respectively and another group of part-Hawaiians scored 29.5. Thus, the group differences found are inconsistent. A group of illiterate Filipino plantation workers from the same 1945 study who had difficulty understanding the test instructions obtained a mean Q-score of 46.38, but this high score was probably more the result of incomprehension and unfamiliarity with pencil-and-paper tasks than of nation group membership.

Two investigations of Australian aborigine Q-score performance (Porteus & David, 1966; Porteus & Gregor, 1963) indicate that these people do, in fact, obtain higher Q-scores than other tested groups. A group of 27 full-blooded and half-caste adolescent aborigines had a mean Q-score of 36. A group of 26 adult male full-blooded aborigines scored 51.69 and a similar group of 23 scored 39.13. Again, these results may have been due to incomprehension and unfamiliarity with the test materials although this is less likely to be true of the adolescent group, who had considerable contact with whites.

Q-Score Conclusion

There is no consistent evidence that members of one ethnic or national group obtain different Q-scores from those of other groups, except in the case of Australian aborigines whose scores are considerably higher than those of other groups, and even this finding may be due to unfamiliarity with paper-and-pencil tasks.

Sex, Age, Social Class, and Education

Test Age

Studies involving sex, age, socio-economic status and education were examined to assess the degree of relationship between these variables and Maze TA. Of these variables, sex produced the most clear-cut and consistent differences in TA. Out of 94 pairs of comparable male and female groups reported in the literature, involving well over 5,000 subjects, the mean TA was higher among the males for 83 of the pairs and higher among females for the remaining 11 (Arthur, 1933a; DuBois, 1944; Fry & Pulleine, 1931; Gruetzow & Brozek, 1946; Jensen, 1966; Joseph & Murray, 1951; Loranger & Misiak, 1960; Louttit, 1931; Louttit & Stackman, 1936; O'Shea, et al., 1942; Phillips, et al., 1948; Porteus, 1919, 1931, 1937, 1955, 1957, 1961; Porteus & Babcock, 1926; Porteus & David, 1966; Porteus & Gregor, 1963; Sparling, 1941; Vernon, 1937; Weisenburg, Roe & McBride, 1936; Zeckel & Van Der Kolk, 1939, Zubin, et al., 1961).

Although there is little doubt of male superiority in Porteus TA among primitive adults and children of all cultures investigated, the superiority of adult males in the Western culture is questionable. Mean TAs for 8 pairs of primitive adult male and female groups have been reported in the literature and in every case the males scored higher than the females (DuBois, 1944; Fry & Pulleine, 1931; Porteus 1931, 1937; Porteus & David, 1966; Porteus & Gregor, 1963). The average difference in score between adult primitive males and females is 2.5 years of TA. Mean TAs for 13 pairs of primitive male and female groups

of children are available and in 12 of these pairs the boys scored higher than the girls (Joseph & Murray, 1951; Porteus, 1937; Sparling, 1941). The exception was among 5 to 7 year-old-Carolinian children from Saipan in the study by Joseph and Murray. In this age group, the girls obtained the higher mean score. Among nonprimitive Western children, 65 pairs of groups have been studied and the boys obtained higher mean TA scores in 58 of these groups (Arthur, 1933a; Louttit, 1931; Louttit & Stackman, 1936; Phillips, et al., 1948; Porteus, 1919, 1955, 1961; Porteus & Babcock, 1926; Zeckel & Van Der Kolk, 1939). Most of the pairs of groups in which the girls obtained higher scores consisted of adolescents. Even among adolescents, however, the male subjects still appear to be superior. Out of 26 pairs of adolescent groups (age 12 - 19) the males obtained higher mean scores in 20 of the pairs of groups. The mean difference in TA between girls and boys amounts to roughly one-half year for subjects under 12 and falls below .2 years for subjects age 12 to 14 when the results of all studies are averaged.

Western adults have not been investigated extensively enough to determine whether or not male maze test superiority is universal. Out of eight pairs of groups, the males scored higher in only five (Gruetzow & Brozek, 1946; Jensen, 1966; Loranger & Misiak, 1960; O'Shea, et al., 1942; Porteus, 1957; Weisenburg, Roe & McBride, 1936; Zubin, et al., 1961).

There is some indication that age, among adults, is negatively related to TA. Out of a total of 18 groups reported in the literature, TA and age were negatively related in 13 (Aaronson,

1957; Frost, 1965; Heston & Cannell, 1941; Jensen, 1966; Loranger & Misiak, 1960; Porteus, Bochner, et al., 1967; Porteus & David, 1966; Shakow & Millard, 1935; Small, 1954; Smith, 1960; Sterne, 1969; Weisenburg, Roe, and McBride, 1936; Zubin, et al., 1961).

Out of 27 groups investigated for the effect of socio-economic status on Maze TA, data indicated that lower class subjects obtain lower TA scores in 21 of these groups (Berry & Porteus, 1920; Havighurst & Janke, 1944; Peterson & Telford, 1930; Porteus, 1919; Weintraub, 1968; Weisenburg, Roe & McBride, 1936). This constitutes reasonably clear evidence of a relationship between social class and TA.

The remaining variable examined was education. Out of nine adult groups investigated, six gave evidence that there is a positive relationship between years of schooling and Porteus TA (Jensen, 1966; Loranger & Misiak, 1960; Shakow & Millard, 1935; Weisenburg, Roe and McBride, 1936; Zubin, et al., 1961). This is not sufficient evidence of a relationship between Porteus TA and education.

Test Age Conclusions

1. Among primitive adults and children of all backgrounds, males obtain higher Maze TAs than do females.
2. Male superiority in Maze performance has not been conclusively demonstrated among Western adults.
3. There is some evidence that age, among adults, may be negatively related to TA. This relationship is only weakly indicated in the literature and requires further investigation.

4. There appears to be a relationship between Maze TA and socio-economic status with individuals of lower social class obtaining lower scores.
5. There is, at the present time, no basis for assuming that a relationship exists between education and TA among adults.

Q-Score

Data are available for the assessment of the relationships between Q-score and sex, age, and social class. Of these relationships, sex and Q-score have been the most extensively investigated.

Nineteen pairs of comparable male and female groups have been reported in the Q-score literature (Foulds, 1951; Porteus, 1942, 1945, 1955, 1961; Purcell, et al., 1962; Sanderson, 1945). In 10 comparisons, females scored higher than males. In the other 9 comparisons, the males obtained the higher Q-scores. Three of these pairs of groups were made up of adult subjects. In 2 of these adult pairs, the females obtained the highest scores, and in the remaining pair, the males scored higher. The remaining 16 pairs of groups were adolescents or children and were evenly split with respect to Q-score; the males obtaining the higher scores for half of the pairs and the females obtaining the higher scores for the other half. The data indicate that there is no sex difference with respect to Q-score, at least among relatively young subjects.

An interesting finding has appeared with respect to the relationship between Q-score and age among children and adolescents.

The subjects on which this finding is based range from about 10 through 17 years. The data indicate that the younger subjects obtain the higher Q-scores. Out of nine comparisons of older with younger normal subjects, the younger subjects scored higher in eight (Porteus, 1955; Sanderson, 1945; Weintraub, 1968). This finding was not substantiated among the delinquents and emotionally disturbed children reported in the literature. Out of five comparisons, the younger group scored higher in only one (Grajales, 1945; Weintraub, 1968). Out of two comparisons among young adults, the younger group scored higher in one (Foulds, 1951; Schalling & Rosén, 1968).

One study (Weintraub, 1968) reported data relevant to the relationship between Q-score and social class. The subjects in this study were divided into three groups on the basis of emotional adjustment. Within each of these groups the lower class subjects obtained higher Q-scores than did the middle class subjects. However, the difference was significant within only one of these groups.

Q-scores of illiterate subjects, totally deprived of formal education, have been found to be quite high (Porteus, 1945; Porteus & Gregor, 1963).

Q-Score Conclusions

1. Q-score appears to be unrelated to sex.
2. For normal subjects, age 10 to 17 years, there appears to be a relationship between Q-score and age such that younger subjects obtain higher scores.

3. The only study pertaining to the relationship between social class and Q-Score suggests a tendency for lower class subjects to obtain higher Q-scores than middle class subjects.

Physical Disorders

Test Age

A few studies have concerned themselves with the effects of physical disease or physical disability upon Maze test performance. Phillips, Berman,and Hanson (1948) compared 52 children who had previously been ill with polio with a group of 52 normal children matched on the basis of age, sex, and intelligence prior to the onset of the disease. The control subjects scored significantly higher on TQ than the polio subjects. However, when subjects were broken down into two age groups, a young group aged 6 to 8 (N= 20) and an older group aged 9 to 18 (N=32), the differences in TQ were significant only for the younger group. Although the mean TQs for all the groups in this study were in the average range, the young control subjects had unusually high TQs (M=135.9). All the other groups scored above average in TQ, but not to such a great extent. Because the significant findings are due to the performance of an apparently atypical group, little weight can be given to the results of this study.

Waite and Neilson (1919) studied the effects of hookworm infestation on children in Australia. Hookworm infestation is reported to produce anemia and toxemia and the authors hypothesized that it would therefore interfere with development. Subjects were

116 children who were free from hookworm, 65 who were classified as having a light infestation, and 159 who were classified as having a heavy infestation. The Maze test was administered in a modified fashion making it relatively easy. No tests of significance were computed in the study which reported that the hookworm-free children generally scored higher than those infested with the disease. The authors assumed that the hookworm infestation produced the retardation in performance and did not consider the possibility that children of low intelligence may be more likely to be hookworm infested.

Three articles report investigations of the Maze test performance of deaf and dumb children (Porteus, 1917, 1918; Zeckel & Van Der Kolk, 1939). Of seven groups of apparently non-retarded deaf children, five were below average in Porteus TA and two gave average performances. The study by Zeckel and Van Der Kolk was the most carefully designed of the three, employing matched controls for each of the deaf groups. In their study, each of the four deaf groups gave poorer performances than their respective control groups, while the control groups generally obtained average scores.

Test Age Conclusion

TA findings in the area of physical disorders are largely inconclusive. There is some suggestion of a deficit in TA among deaf and dumb children.

Q-score

No Q-score studies have been reported in which the performance of normal subjects and physically diseased or disabled subjects can be compared.

Chlorpromazine, Other Drugs and Vitamin B

Test Age

There are eight studies in the literature investigating the effects of chlorpromazine on Maze performance. At first glance, these studies seem equally distributed with four reporting a decrement in Maze performance under chlorpromazine (Helper, et al., 1963; Porteus, 1957; Porteus & Barclay, 1957; Porteus & Ching, 1959), and four reporting no drug effect on Maze score (Aaronson, 1963; Grygier & Waters, 1958; Judson & MacCasland, 1960; Petrie & LeBeau, 1956). There are, however, flaws in three of the studies giving negative results that would tend to cast doubt on the validity of their findings, and the remaining negative study reported group differences in the expected direction which failed to reach significance.

Petrie and LeBeau (1956) made no attempt to control for practice either by using different forms of the Maze test or by using a control group. Aaronson (1963) used no control group and it is unlikely that he used alternate forms of the test. In both of these studies, the Maze test was reapplied after a relatively short period of drug administration. Petrie and LeBeau (1956) tested after 9 days on the drug and Aaronson after 2 weeks. While 2 weeks may be barely adequate, 9 days is generally considered insufficient for studies of this type of drug. Porteus (1965) recommends 6 weeks.

The only fault that casts doubt on the findings in the otherwise carefully executed study by Grygier and Waters (1958) is that the nature of the maze test employed is not at all clear. The

authors state that they used "a simple perceptual maze test" that was "an adaptation of the Porteus Mazes". There is no way to determine how this inadequately described adaptation resembles the original.

Of the five groups of subjects involved in studies where practice was controlled, and the Mazes were administered in unmodified form, all five showed a decrease in mean TA following chlorpromazine administration. The mean of the reported decreases amounts to approximately 1-1/2 years of TA.

Five other studies of drug effect on the Mazes have been done, each investigating a different drug: butyrophenones, LSD-25, methylphenidate (Ritalin), dexedrine and L-dopa. No differential effect was found between two different butyrophenone compounds in a study by Aaronson (1963). Aronson and Klee (1960) found significant differences in TA two hours following drug administration between a control group, a group given 75 mcg of LSD and a group given 100 mcg of LSD. Their scores were 16.27 (N=34), 15.21 (N=28) and 14.25 (N=6) years of TA respectively. Conners and Eisenberg (1963) found a significant difference in TA between a group of emotionally disturbed children given Ritalin for 10 days and a comparable control group given a placebo for the same period of time, suggesting that Ritalin has a facilitating effect on Maze performance among subjects of this type. The facilitating effect appeared to be greatest for children of lower intelligence. No significant difference was observed between the experimental and control groups of high intelligence. Conners and Eisenberg, in a personal communication with Porteus

(Porteus, 1965), also report a facilitating effect on Maze performance associated with dexedrine. A preliminary report (Meier & Martin, 1970) suggests that the drug L-dopa may have a facilitating effect on the Maze performance of parkinsonian patients.

Two studies have attempted to demonstrate the effects of Vitamin B deficiency upon Maze performance (Gruetzow & Brozek, 1946; O'Shea, et al., 1942). O'Shea et al., compared four experimental subjects deprived of Vitamin B and four controls. An idiosyncratic method of scoring the Mazes was used. Several other intellectual and motor tests were also administered. It was reported that Maze scores were impaired by Vitamin B complex deficiency and improved after therapy; however, none of the remaining test scores were affected by the Vitamin deficiency or therapy. Unfortunately, the results were confused by the undefined Porteus scoring, the extremely small sample size, and inadequate controls for practice effects on the Mazes.

Gruetzow andBrozek (1946) attempted to replicate these findings with eight subjects and found no significant changes in Maze scores. However, they, too, used an idiosyncratic scoring method and failed to take practice effects into account.

At present there is no convincing evidence that Vitamin B complex deficiency affects Maze performance, but a well designed study might well demonstrate such effects.

Leibowitz (1966) reports the Maze scores of drug addicts, but his findings do not necessarily relate to the effects of the drugs themselves.

Test Age Conclusions

1. Evidence published to date suggests that chlorpromazine produces a deficit in TA of about 1-1/2 years. Studies failing to demonstrate this effect were of faulty design. Further research is needed before this conclusion can be considered established.
2. It would appear that drugs other than chlorpromazine may affect Maze performance. LSD was found to interfere with Maze performance, and Ritalin was found to facilitate Maze performance, at least in emotionally disturbed children of low intelligence.
3. What little research has been done with Vitamin B and drug addiction is inconclusive.
4. A great deal more work needs to be done in the area of drug effect on maze TA before test results of subjects taking drugs can be interpreted with any confidence.

Q-Score

Two studies report Q-score results with chlorpromazine. Helper et al., (1963) report a decline in Q-score associated with the administration of chlorpromazine falling just short of significance ($p=.07$), and a significant increase in Q-score following chlorpromazine removal ($p=.05$). Aaronson (1963) reports a nearly significant Q-score reduction in one of his experimental groups with the introduction of chlorpromazine, and a significant overall reduction for both groups combined.

While the results of the study by Helper, et al., (1963) are suggestive of an improvement in Q-score associated with chlorpromazine, Aaronson's study is complicated by the possibility that

he failed to control for practice. Since Q-score tends to decrease with repeated application of the Maze test, it is possible that the observed change in Q-score was due to practice rather than the effect of the drug. Aaronson (1963) also investigated the differential effect of two butyrophenones on Q-score and found no difference.

Two studies (Leibowitz, 1966; Schalling & Rosén, 1968) report Q-scores of alcoholics and drug addicts, but their findings do not necessarily relate to the effects of the drugs themselves.

Q-Score Conclusion

Sufficient evidence of a drug effect on Q-score is available to warrant further investigation into the existence and nature of the effect.

Chapter V

Interpretation and Validity

Interpretation and Validity

Planning Capacity

Test Age

The Maze test was originally developed to test planning capacity of retarded children. Since the test's introduction in 1915, surprisingly little work has been done to evaluate its validity as a measure of planning ability. Probably the chief reason so little work has been done is that it is very difficult to set up an adequate criterion of planning ability against which to evaluate the Maze test.

A second reason is that the test itself is unquestionably face valid. It is necessary, in tracing through a maze, to decide in advance which direction to push the pencil at each choice point. To do this successfully, the subject must survey the entire maze, identify the various alternative routes open to him, evaluate the alternatives, and act according to his evaluation. It would appear to be virtually impossible to successfully work through the entire series of mazes in any other manner, and it is probably generally agreed that this procedure closely approximates "planning".

There are only two studies in the literature that attempt to validate the Maze test as a measure of planning ability (Porteus, 1920; Small, 1954). The subjects in the study by Porteus were 38 male and 44 female institutionalized retardates, age 15 to 30 years. These subjects were rated on a 4-point scale for lack of planning capacity by the school department principal at the institution (familiar to readers of the early

Maze literature as Vineland's ubiquitous Mrs. Nash). Among the male subjects, planning capacity correlated .66 with Porteus TA, and among the female subjects the correlation was .75. For the purposes of this study, planning capacity was defined as the ability to master a new task or job.

These same subjects were also rated for a trait labeled "irresolute", defined as the lack of ability to make independent decisions. This trait, as defined in this study, sounds as though it would be closely related to planning ability. The correlation between irresolute and Maze TA among the males was .59 and among the females it was .69. This study indicates that traits that appear to be related to planning ability produce highly significant correlations with Porteus Maze TA.

Porteus also found highly significant correlations between these two traits and Binet MA in the same two groups of subjects. The actual values of the correlations between the Binet and planning capacity were conspicuously lacking, but these values were reported for "irresolute" as .69 and .74 for the male and female subjects respectively. This finding argues against Porteus' claims that his test measures planning capacity while the Binet does not.

Small (1954) investigated the relationship between Maze TA and three measures of planning, independent of intelligence test scores, in a group of 162 nonpsychiatric Veterans Administration Hospital patients during the recovery period of their illness. The measures of planning used were three self-administered scales designed for this study. One scale was made up of MMPI items considered to reflect planfulness (scale AB); the second scale contained statements about everyday activities which were felt to be related to

planning (scale PS); and on the third scale, subjects were asked to rate themselves on a 6-point basis for the amount of planning done in everyday life (Self Rating Scale). Superficially, these scales would appear to be more closely related to the extent that subjects are in the habit of making plans rather than to the extent that they are capable of making plans. The correlations between Maze TA and the three planning scales, uncorrected for intelligence, were -.22, .38, and -.28 respectively. Because scale AB and the self-rating scale are negatively related to planning, these correlations indicate a positive relationship between the Porteus and planning; significant, in each case, beyond the .01 level of confidence. The three planning scales were also significantly correlated with each other. Partialling out verbal and nonverbal IQ (as measured by the SRA verbal and nonverbal forms) had virtually no effect on the correlations between Maze TA and the three planning scales indicating that the Mazes were related to planning independent of intelligence in this study.

Test Age Conclusions

1. These two studies, viewed conjointly, suggest that the Porteus Mazes are measures of planning to a certain extent, even with intelligence ruled out.
2. Two studies are inadequate to test the validity of claims that have been repeatedly made by Porteus and accepted by psychologists for the past 55 years. A great deal more work is needed in this area. It would be particularly valuable to devise additional, perhaps more meaningful, criteria against which to validate the test.

Intelligence

Test Age

By far, the intelligence test most extensively correlated with Maze TA has been the Binet. Fifty-two separate correlations involving a total of over 3,000 subjects have been reported in the literature, (Arnold, 1951; Brill, 1937; Cornell & Lowden, 1923; Cunningham, 1916; Earle & Milner, 1929; Ebert & Simmons, 1943; Gaw, 1925b; Louttit & Stackman, 1936; Michaels & Schilling, 1936a; Morgenthau, 1922; Porteus, 1918, 1919, 1923; Porteus, Dewey & Bernreuter, 1930; Porteus & Kepner, 1944; Ross, 1921; Satter & McGee, 1954; Shakow & Millard, 1935; Sparling, 1941; Vernon, 1937; Weisenberg, et al., 1936; Worthington, 1926). These correlations range from +.10 to +.83. In eight of the reported correlations, no correction was made for the age of the subjects. Since both the Binet MA and Maze TA are strongly related to age in young subjects, failure to correct for age artificially inflates the correlation between these two tests. In general, the wider the range in age of the subjects, the greater the artificial inflation in correlation. The correlations not corrected for age range from .59 to .83 with a mean of .73. Two studies (Morgenthau, 1922; Ross, 1921) did not report whether or not age was corrected in their correlations.

The remaining correlations between the Binet and the Maze test were all corrected for age by using subjects over 14, by correlating IQ with TQ or by using subjects all of the same age. These 42 age-corrected correlations involve over 2,000 subjects and range from +.10 to +.75. The mean of all the published age-corrected correlations was +.50 (using the Z Score conversion method of calculation).

The correlations with the Binet seem to be fairly typical of most of the correlations found in the literature between intelligence tests and the Maze test. Four separate correlations with the Mazes have been reported for the Raven's Progressive Matrices, the WAIS Performance IQ, six for the WAIS or Wechsler Full Scale IQ, and four have been reported for the Goodenough Draw-A-Man (Docter, 1960b; Gibbens, 1958; Leiter, 1959; Loranger & Misiak, 1959, 1960; Peterson & Telford, 1930; Satter, 1955; Satter & McGee, 1954; Sterne, 1969; Tizard, et al., 1950; Tobias & Gorelick, 1962; Weisenburg, et al., 1936). All these correlations have been within the range reported for the Binet, and the means of reported correlations for each of these tests do not differ significantly from .50.

Individual correlations for a number of other measures of intelligence have been reported which are also consistent with the results reported for the Binet (Bennett, 1956; David & Bochner, 1967b; Dentler & Mackler, 1962; Earle & Milner, 1929; Gibbens, 1958; Gregor & McPherson, 1963; Leiter, 1959; Loranger & Misiak, 1959, 1960; Peterson & Telford, 1930; Porteus, 1919; Porteus, Dewey & Bernreuter, 1930; Shakow & Millard, 1935; Small, 1954; Tizard, et al., 1950; Tobias & Gorelick, 1962; Weisenburg, et al., 1936). Only atypical findings will be discussed.

Gaw (1925b) reported correlations between Maze TA and each of a battery of 13 nonverbal IQ tests for 100 school children, age 13 years. Separate correlations were reported by sex. Of 26 correlations, 22 were below +.20 and 4 of these were negative. Compared to other studies, Gaw's results appear to be unusual. She used,

among other tests, formboard tests, object assembly tests, Knox cubes, the Healy Picture Completion II, a digit symbol test, and a block design test. These tests or very similar tests have been correlated with the Mazes in other studies (Earle & Milner, 1929; Loranger & Misiak, 1959, 1960; Peterson & Telford, 1930; Porteus, 1919; Porteus, Dewey & Bernreuter, 1930; Roberts & Alcuaz, unpublished data; Satter, 1955; Shakow & Millard, 1935; Tizard, et al., 1950; Tobias & Gorelick, 1962), and of the 28 reported correlations, only one (Earle & Milner, 1929) was below +.20. It is difficult to imagine why Gaw's correlations were so consistently low, especially when the correlations she reported between the Binet and the Maze test were comparable to other reported correlations between these two tests.

Earle and Milner (1929) replicated Gaw's study using the Mazes and five of the other tests used by Gaw with a comparable group of subjects, London school children of age 13. These investigators found correlations between the Maze test and other tests that were substantially higher than Gaw's. They reported a correlation of .14 between a formboard test and Maze TA among their 137 female subjects. The remaining 9 correlations for the female and 113 male subjects ranged from .23 to .39 which is more in accord with results from other studies.

Arnold (1951) reported some unusually high correlations between the Binet, a modification of the Leiter International Performance Scale and a modification of the Maze test, ranging from .78 to .89 for three groups of subjects. Although her correlations were not corrected for age, it was possible to make an age correction because the author included a table of individual

subject IQs on each test. This resulted in a drastic reduction to unusually low Binet-Maze correlations of .02 for a group of 25 feeble-minded children, the lowest ever reported between these two tests, and .12 for 25 normal children. The Binet was not administered to the third group, 25 cerebral palsied children, because they were unable to take it. The age-corrected correlations between Maze and Leiter IQ were .48, .33, and .82 for the feeble-minded, normal, and cerebral palsied groups, respectively. It is difficult to resolve the differences between the raw data and the reported correlations in this study; therefore, the findings cannot be given much weight.

Another unusually low correlation between Maze score and IQ was found in a study by Erikson and Roberts (1971). These investigators used scores from three different IQ tests, the WISC, the WAIS, and the Binet for different subgroups of their sample of 40 delinquent boys and obtained a correlation of .03 between IQ and Maze TQ. Two other studies are available that used similar combinations of different IQ tests with different subjects and reported the correlation of the resultant IQ scores with the Mazes. Docter (1960a) found a correlation of .11 between the Maze test and IQs from the WISC, California Test of Mental Maturity, and the Wechsler-Bellevue I, which is also quite low. Weintraub (1968), however, reported a correlation of .41 between Maze TQ and IQs from the WISC, the Binet and the Lorge-Thorndike. It is difficult to rationalize these findings.

Of all types of intelligence tests, the most consistently low correlations with the Porteus are obtained with vocabulary

tests. Correlations between the Porteus and various vocabulary tests for five groups made up of a total of 411 subjects ranged from -.07 to +.34 with a mean r of +.13 (Bennett, 1956; Roberts & Alcuaz, unpublished data; Tizard, et al., 1950; Weisenberg, et al., 1936). It may well be that the Maze test and tests of vocabulary are measuring two different aspects of intelligence. A factor analytic study could prove valuable on this point.

Verbal intelligence batteries, such as the WAIS verbal section, appear to fall between vocabulary tests and other tests of intelligence in their degree of relationship with the Porteus. There are 16 correlations reported in the literature between verbal intelligence and TA involving 1806 subjects (Dentler & Mackler, 1962; Gibbens, 1958; Jensen, 1966; Leiter, 1959; Roberts & Alcuaz, unpublished data; Satter, 1955; Satter & Mc Gee, 1954; Small, 1954; Tobias & Gorelick, 1962; Weisenburg, et al., 1936). The mean of the reported correlations is .29 which, again, is significantly less than the correlation of .50 found with the Binet.

Rather clear-cut evidence, of a different sort, that Maze TA is related to intelligence comes from the reported test scores of 13 groups of institutionalized retardates, all of which are well below the score that would be expected for normal subjects (Arnold, 1951; Brill, 1937; Cooper, et al., 1967; Doll, et al., 1932; Herd, 1923, 1930; Michaels & Schilling, 1936a; Porteus, 1917; Satter & Mc Gee, 1954; Tizard, et al., 1950; Tobias & Gorelick, 1962). These studies indicate that subjects of low intelligence, as operationally defined by institutionalization for mental retardation, obtain low Maze TAs.

Achievement and learning can be seen as related to intelligence. The correlational findings with various measures of achievement (teacher ratings of achievement, class standing, and tests) closely parallel the findings with intelligence tests. Twenty-five reported correlations for 20 groups of subjects between measures of achievement and Maze scores ranged from .18 to .60 with a mean of .39 (Berry & Porteus, 1920; David & Bochner, 1967b; Frost, 1965; Porteus, 1919; Tizard, et al., 1950; Weisenburg, et al., 1936). Correlations between the Mazes and tests of learning are quite a bit lower, ranging, for 20 correlations, from -.02 to +.46 with a mean of .28 (Helper, et al., 1963; Jensen, 1960, 1966; Morgenthau, 1922). Eliminating correlations that may have been inflated due to the influence of age, the range becomes -.02 to +.27 and the mean, .16. Most of the reported correlations are with Jensen's alternation learning in which subjects must discover through trial and error, the proper sequence for a series of switches and turn them on in order. Other correlations are with rote learning tasks such as paired associate learning.

Test Age Conclusions

1. Correlations between the Maze test and intelligence tests tend to fall in the .3 to .6 range.
2. Failure to correct for age in young subjects tends to inflate the correlations.
3. Correlations between vocabulary tests and Porteus scores suggest that these measures are relatively independent of TA.
4. Measures of achievement appear to be similar to tests of

intelligence in their relationship with the Mazes.

5. Learning ability, of the types investigated to date, does not appear to be highly related to Maze ability.

Q-score

Of 13 reported correlations between Q-score and various tests of intelligence, 5 were significant beyond the .05 level of confidence (3 of these were beyond the .01 level), but in no case did the scores on intelligence tests account for more than 18% of the total variance in Q-score (Bennett, 1956; Docter, 1960a; Docter & Winder, 1954; Erikson & Roberts, 1971; Gibbens, 1958; Grajales, 1945; Porteus, 1942; Roberts & Alcuaz, unpublished; Schalling & Rosén, 1968; Weintraub, 1968; Wright, 1944). The highest correlation reported was -.42 with the Raven's matrices for a group of institutionalized delinquents. The lowest correlations were .0 with the Shipley-Hartford for a group of hospital patients (Bennett, 1956), and, in another study (Erikson & Roberts, 1971), .0 with the Binet, WAIS or WISC for a group of institutionalized delinquents. The only reported positive correlation was found by Bennett (1956) with Shipley Vocabulary. The mean correlation over all the available studies is -.21, involving just over a thousand subjects. Indications are that there is a low order negative correlation between intelligence test scores and Q-score. This correlation, because it is small, fails to reach significance in many studies, particularly in studies with a small sample size.

In a study of 191 institutionalized delinquents, Gibbens (1958) found a correlation of -.20 with a test he labeled

"Spatial test." Unfortunately the nature of this test was not described in the article, although it was stated that the test is a measure of intelligence. The results with Gibben's Spatial test were not used in calculating the total correlation given above. If the Spatial test does in fact measure intelligence, the results are in close agreement with the results for other intelligence tests.

Two additional articles mentioned Q-score correlations, but did not give exact figures. Craft (1965), in a study involving 67 subjects, stated, "the high 'Q' score correlated to a highly significant extent with low intelligence..." He did not say what the correlation was. Helper, et al., (1963) reported that the correlation between Q-score and each of two rote learning tasks (and another measure not relevant to this discussion) was .11 or below for 39 subjects, but he did not specify which measure gave the highest correlation nor what the other correlations were.

Q-Score Conclusion

There is a low order negative correlation between Q-score and intelligence test scores, indicating that Q-score is largely, but not completely, independent of intelligence.

Brain Damage

Test Age

Pre- and postoperative Maze test results have been reported for operations involving the frontal lobes, the temporal lobes and portions of the diencephalon. In addition, there have been some investigations of the Maze test's relationship to other

criteria of brain damage, such as psychiatric diagnosis, neurologic signs and psychological tests. Most of the studies relevant to the effect of brain damage on TA deal with surgically induced damage to the frontal lobes in psychiatric patients and in a few cases of intractable pain (Crown, 1952; Jones, 1949; King, 1949; Landis, 1951; Landis, et al., 1949; Porteus & Kepner, 1944; Porteus & Peters, 1947; Robinson, 1946, 1950; Sheer, 1956; Sheer, et al., 1952; Smith, 1960; Koskoff, et al., 1948; Landis & Erlick, 1950; Landis et al., 1950; Malmo, 1948; Medina, et al., 1954; Petrie, 1952; Petrie & LeBeau, 1956; Smith & Kinder, 1959; Tow, 1955; Wittenborn & Mettler, 1951).

In general, the studies indicate that frontal lobe damage results in decreased Maze ability. Losses in TA at the first postoperative testing are reported for 25 of the 28 groups in these studies. The difference in TA before and after psychosurgery is reported for 26 groups involving a total of 335 subjects. The mean postoperative loss for all these groups is 1.5 years of TA. The test results reported in these studies are complicated by several factors, and the above analysis is a somewhat unrealistic over-simplification. The important factors that must be taken into consideration are: The effect of practice on Maze TA, the length of the interval between the operation and the postoperative testing, the locus of the operation within the frontal lobes and the mean preoperative TA of the subjects.

The effect of practice on TA has previously been discussed, and it was concluded that TA could be expected to increase by 1.25 years on the second administration of the test and by an additional

0.7 years on the third. Adding the appropriate practice correction to the group losses reported in each study, the mean practice corrected postoperative loss in TA becomes 2.3 years, and with three studies of somewhat questionable design (Jones, 1949; Robinson, 1946, 1950) eliminated from the calculations, the mean practice corrected loss is 2.5 years.

The effect of the length of the postoperative interval on loss of TA can be examined through the analysis of studies differing in the length of time elapsed between the operation and the first postoperative application of the Maze test. The most extensively investigated postoperative intervals are 10 days, 3 to 4 weeks and 3 months (Crown, 1952; King, 1949; Petrie, 1952; Porteus & Kepner, 1944; Porteus & Peters, 1947; Sheer, 1956; Sheer et al., 1952; Wittenborn & Mettler, 1951). Averaging the results for all these studies, the mean losses in TA at the first postoperative testing for the three intervals, corrected for practice, are 3.4 (N=26), 2.7 (N=43), and 2.3 (N=85) years respectively. On the basis of these data, it appears that there is an initial sharp loss in Maze ability followed by a negatively accelerated partial gain which fails, at least in the first 3 months, to bring the subjects back up to their preoperative level.

Further conformation that there is an initial postoperative decline in TA followed by a partial recovery can be found in the 1947 study by Porteus and Peters. The first postoperative test was administered to the subjects at irregular intervals owing to lack of space and personnel. The authors reported that 16 subjects tested within one month after surgery lost 2.16 years of TA;

21 subjects tested between 1 and 1-1/2 months after surgery lost 2.10 years, and 18 subjects tested more than 1-1/2 months after surgery lost 0.67 years.

Some investigators reported data relating to the site of the frontal lobe operation as well as the length of the post-operative interval (King, 1949; Petrie, 1952; Sheer, et al., 1952). The surgery investigated in these studies can be divided into anterior and posterior frontal lobe operations (see Figure 1). Only surgery explicitly involving these widely separated areas will be included in the analysis presented here to avoid, as much

---

Insert Figure 1 about here

---

as possible, overlap due to surgical error. Portions of the frontal lobes between these two areas were undoubtedly included in both the anterior and the posterior operations. The first postoperative testing was done 10 days, 3 to 4 weeks, and 3 months following frontal lobe surgery in these studies for both the anterior and the posterior operations. The mean losses in years of TA, corrected for practice, for the anterior operations were 2.7 (N=14), 1.3 (N=13) and 0.9 (N=8) at the three postoperative intervals respectively. These losses were 4.6 (N=10), 3.7 (N=11), and 3.1 (N=32) for the posterior operations. These data indicate that the effect of the postoperative interval on losses in TA noted above applies to both types of frontal lobe operations, but the effect of surgery on TA is more severe following posterior operations.

Additional information on the relationship between TA and the area of the frontal lobes operated can be found in Crown (1952), Petrie (1952), Smith (1960) and Smith and Kinder (1959). Crown

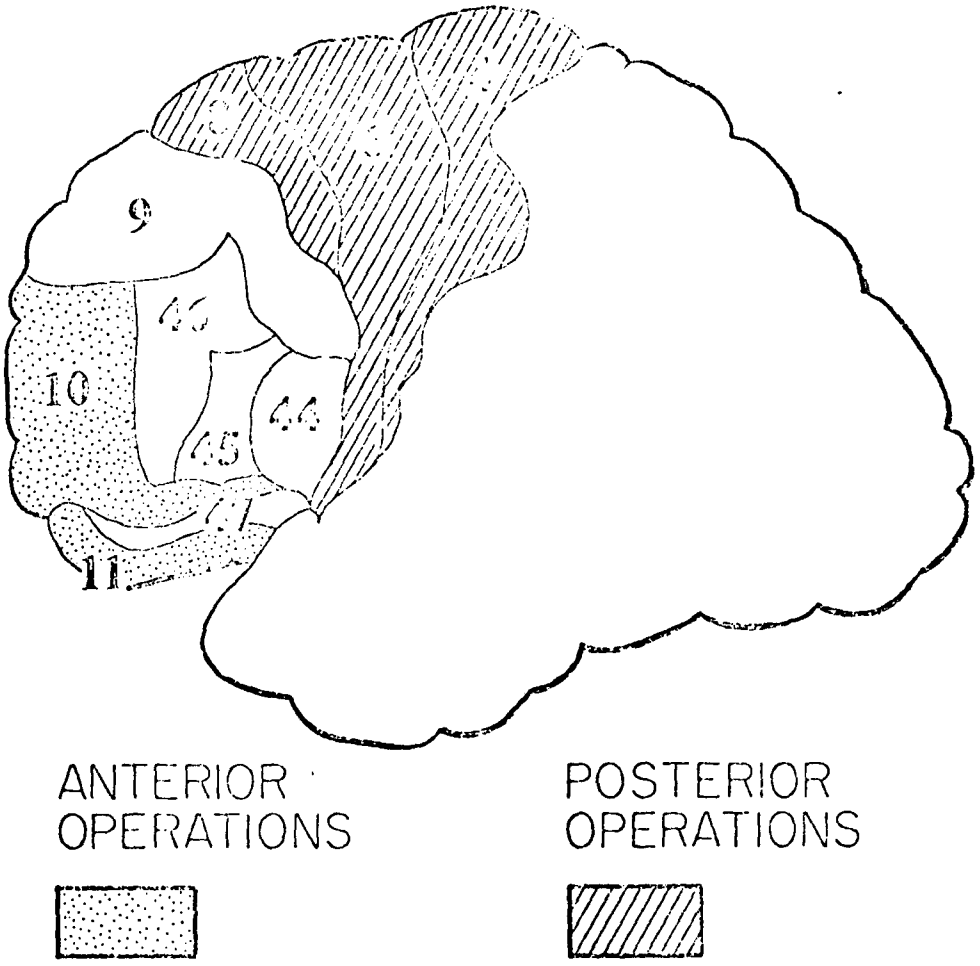


Fig. 1. Lateral view of the left hemisphere of the brain showing the site of anterior and posterior operations with respect to Brodmann's areas.

(1952) reported that losses in TA were significantly greater following lobotomies performed by a surgeon said to have operated in a plane posterior to that of two other surgeons when compared to the losses following operations by the other two surgeons.

Petrie (1952) reported a mean loss in TA for 15 subjects 6 months after an anterior frontal lobe operation which, when corrected for practice, amounted to 1.3 years. This finding is quite consistent with the pattern of postoperative TA losses found for shorter postoperative intervals with anterior frontal lobe operations.

The two articles by Smith (Smith, 1960; Smith & Kinder, 1959) concern a follow-up study done 8 years after frontal lobotomies were performed. The group from this study with the more posterior operation showed a loss of 3.9 years on the 8 years postlobotomy test as compared to the first prelobotomy test while the anterior group lost 0.2 years of TA. An unoperated control group lost 0.65 years of TA over the 8 year period, apparently due to aging. Again, these results generally agree with results from other studies investigating the differential effects of anterior and posterior lobotomies.

The remaining factor that could have influenced the Maze test results following frontal lobe surgery is preoperative TA. Unfortunately, there are few data available that relate to this question and are not contaminated by other factors which appear to be related to postoperative losses in TA. Four pairs of groups differing in preoperative mean TA, but matched for other variables, are available for comparison (Sheer, 1956; Sheer,

et al., 1952). For two of these pairs of groups, the group showing the greater postoperative loss in TA obtained the lower preoperative score. For one pair, the group showing the greater postoperative loss in TA obtained the higher preoperative score. For the remaining pair, the postoperative losses were the same while the preoperative TAs differed by 4.5 years. On the basis of this evidence there does not appear to be any relationship between preoperative TA and postoperative loss of TA.

The permanence of the loss in TA following frontal lobe surgery has been of interest to many investigators. Due to faulty experimental design, early investigators concluded that the loss in Maze ability lasted only a few months postoperatively (King, 1949; Landis, 1951; Landis, et al., 1949; Landis, et al., 1950; Petrie, 1952; Sheer, et al., 1952). These investigators all failed to take practice effects into account. The magnitude of practice gains could quite easily have offset the postoperative loss in TA. Later studies, more sophisticated in design, of which Sheer (1956) is an outstanding example, found the loss in Maze test ability to last for at least 3 months. Only two long term follow-up studies have been published (Medina, et al., 1954; Smith, 1960). Smith's (1960) study indicated a permanent decrement in Maze test ability following more posterior frontal lobe operations. An 8 year postoperative follow-up study by Medina, et al., (1954) failed to give unequivocal results. The authors reported a loss in Maze score, but were unclear as to whether it was a loss in TA or Q-score. Although it is possible that losses in Maze ability following anterior frontal lobe

operations are recovered over extended periods of time, to date there has been no evidence offered that losses following posterior frontal lobe operations are ever recovered.

Two other interesting Maze test findings emerge from the literature of frontal lobe operations. The first is a tendency for subjects to increase the number of repeated errors made (i.e., entrances two or more times in succession into the same blind alley) following surgery. Petrie (1952), Porteus and Kepner (1944), Porteus and Peters (1947) and Tow (1955) all reported increases in repeated errors following frontal lobe surgery. Two of these authors (Petrie, 1952; Tow, 1955) reported that the increase reached significance, while the significance of the increase was not tested in the two other studies. Landis and Erlick (1950) reported no postoperative increase in repeated errors.

Three studies (King, 1949; Landis, et al., 1949; Porteus & Peters, 1947) reported that a "drop and rise" pattern of Maze scores was associated with social recovery among their subjects. A "drop and rise" pattern is characterized by a large initial postoperative loss in TA followed by a marked improvement in score with repeated postoperative testing.

Two studies have been done to test the effect on Porteus TA of the unilateral removal of temporal lobe tissue in patients suffering from temporal lobe epilepsy (Brown, et al., 1956; Meier & French, 1966). The study by Brown, et al., involved a total of six subjects, three of whom were tested with the Porteus before and after surgery. These three subjects showed a significant postoperative loss of TA amounting to 4.4 years, after correcting

for practice. All three of the subjects had right hemisphere operations.

The study by Meier and French (1966) was much more extensive involving a total of 65 subjects subdivided into four groups according to the hemisphere operated and the time of postoperative testing. Practice was controlled by using the Vineland Revision preoperatively and the Extension Series postoperatively. No postoperative change in score was observed in any of the groups tested.

It is probably safe to conclude that postoperative changes in TA following temporal lobe surgery have not been adequately demonstrated. The available evidence points to the conclusion that such changes in TA do not occur, at least not in epileptic subjects. It is likely that this finding applies to normals as well. None of the groups of subjects was markedly subnormal in preoperative Porteus TA, the lowest scoring group having a mean TA of 12.97. This means that the subjects were essentially normal with respect to TA prior to the operation and could reasonably be expected to reflect the effect of extensive temporal lobe damage on Porteus Maze ability.

To date, the greatest mean loss in Porteus Maze TA following psychosurgery was reported by Sheer, et al., (1952) for two patients who underwent bilateral thalamotomies. These subjects lost 4.2 years of TA. Adding a practice correction of .7 (the subjects were given two preoperative tests), this loss becomes 4.9 years. In spite of the magnitude of the postoperative change in score, this finding cannot be given much weight because of the extremely small N involved.

Additional evidence that operations in the area of the

thalamus result in decreased TA comes from a study by Meier and Story (1967) in which subjects with Parkinson's disease were given right or left subthalamotomies. The 17 left subthalamotomy subjects had a non-significant postoperative loss of .85 years of TA. Practice was controlled by use of the Extension Series postoperatively. The group of 12 subjects who underwent the right-sided surgery, however, showed a significant loss of 3.21 years. This is the first published evidence of any differential effect of laterality on Maze losses following injury to the brain.

One other investigator, Petrie (1952), briefly commented on a study of the differential effects of right and left hemisphere surgery. Two groups of four subjects each underwent anterior frontal lobotomies. One group had left-sided surgery and the other group had right-sided surgery. Each group sustained slight postoperative decreases in TA, but did not differ from each other in this respect. Meier and Okayama (1969) found no difference in TA between 117 right and left hemisphere stroke victims. These studies suggest that there is no laterality at cortical levels with respect to Porteus Maze ability. The question of the effect of laterality on Porteus TA has been almost entirely neglected.

There are a number of other studies that appear to be related to the question of the effect of brain damage on TA (Aaronson, 1957; Bennett, 1956; Doll, 1932; Good, 1957; Meier & Okayama, 1969; Michaels & Schilling, 1936a; Myatt, 1951; Satter, 1955; Sterne, 1969; Sydow, 1953; Weisenburg & McBride, 1935). These studies all involve the relationship of TA to some criterion of

brain damage such as psychiatric diagnosis, neurologic signs or psychological tests purported to measure brain damage. The results are inconsistent. The Maze test appears to be related to the criterion in some studies, but not in others. This is not surprising considering that Maze test ability appears to be localized to an area of the brain that is generally neurologically silent. The inconsistencies can probably best be explained as differences among the various criteria in their relationship to brain damage in general and frontal lobe damage in particular. Because of these considerations, the bulk of the studies mentioned provide little insight into the relationship of TA to brain injury.

Some interesting findings were presented in the article by Meier and Okayama (1969) who investigated stroke victims in Minnesota and Japan. These investigators found a significant difference in TA between the 33 Japanese and 84 American stroke patients studied, the Japanese average approximating that of the general population of normals and the American average showing considerable deficiency. No difference in age or the locus of cerebral involvement could be found that would account for the observed difference in scores. On the basis of these results it must be assumed that either the Japanese subjects were superior in TA to the Americans prior to the stroke or that the Americans suffered greater losses as a result of the stroke

The only cross cultural Porteus Maze data available comparing Japanese with American subjects (Porteus, 1937, 1939a, 1950, 1961; Porteus & Babcock, 1926) are based on the test results of children and do not indicate differences in performance in favor of either nationality. Only one study has been published with

information relating to the possibility of a cross cultural difference in the effect of brain damage on Porteus TA. Individual results for five Japanese frontal lobotomy subjects were reported by Porteus and Kepner (1944). Correcting their data for practice by subtracting 1.25 years from each postoperative score, four of the subjects appear to have definitely declined in TA. The fifth subject was reported to have had an inadequate operation according to postoperative X-rays. No other information on the reaction of Japanese subjects to cerebral injury is available. This evidence, albeit scanty, does not indicate a difference in response to frontal lobe injury among the Japanese. Thus, the available evidence argues against both possible interpretations of Meier and Okayama's results and the explanation for their findings will have to await further investigation. In a pilot study by Meier reported in this same article, it was found that when the Porteus was administered to American subjects approximately 2 days after hospitalization for a stroke, TA was positively related to subsequent neurological improvement assessed 10 days and 1 year following the stroke.

Good, (1957) reported results for the only available study of the effect of shock treatment on TA. Unfortunately, practice was not adequately controlled in this study, and the results were further confounded by inconsistent use of the Extension Series. Correcting for practice as best as possible, it appears that a group of 18 subjects judged "improved" following a course of regressive shock treatment showed no change in TA on a 2 to 3 week post-shock testing or on a 3 month to 1 year follow-up testing, while a group of 9 "unimproved" subjects lost about

a year of TA on the immediate post-shock testing and lost an additional half year on the follow-up testing.

Test Age Conclusions

1. There is a postoperative loss in TA within the first 3 months following frontal lobe surgery, which can be estimated to be approximately 2 to 2-1/2 years with practice ruled out.
2. It appears that there may be a sharp immediate postoperative loss in TA followed by partial, but by no means complete, recovery within 3 months postoperatively.
3. There is a greater loss in TA associated with more posterior frontal lobe operations than with more anterior ones.
4. Evidence of recovery of the TA lost following posterior frontal lobe operations has not been reported.
5. There is some evidence that repeated errors are more likely to occur following frontal lobe surgery.
6. There appears to be some indication that TA losses are associated with damage to the thalamus and subthalamus, with a possibility that these losses are associated with right-sided, rather than left-sided, damage. Further investigation is needed in this area.
7. It is an open question as to which other areas of the brain are associated with Maze ability. There has been no convincing demonstration of losses in TA following damage to areas of the brain other than the frontal lobes and certain portions of the diencephalon. However, the

research has been inadequate, the temporal lobes being the only other area specifically studied.

8. The Maze test shows sufficient promise as a diagnostic instrument, and perhaps even as a prognostic instrument, in evaluating certain types of brain damage, to warrant further investigation. Inquiry should be directed to the Maze test's sensitivity to areas of the brain other than the frontal lobes, to the relationship of initial TA to subsequent losses and to the relationship of immediate post insult TA to recovery from the incapacitating effects of brain damage.
9. The effect of shock treatment on TA should be further investigated.

Q-Score

There are seven studies relating to the effect of brain surgery on Q-score (Brown, et al., 1956; Crown, 1952; King, 1949; Landis & Erlick, 1950; Medina, et al., 1954; Porteus & Peters, 1947; Sheer, et al., 1952). Four of these studies reported mean pre- and postoperative scores. There were no significant postoperative changes in Q-score following either frontal or temporal lobe operations in these four studies.

King (1949) reported that the analysis of the Q-scores of subjects given frontal lobe operations and of unoperated controls produced no significant results. Medina, et al., (1954) conducted an 8 year postoperative follow-up study of postoperative frontal lobotomy effects and reported that their operated subjects gave an inferior Maze performance to their control subjects, significant

at tne .01 level. The authors were unclear as to whether this was a difference in Q-score or TA. Sheer, et al., (1952) reported a slight increase postoperatively in the number of pencil lifts per Maze made by their subjects. The subjects were subdivided into five groups according to the type of operation performed. Four of the groups had frontal lobe operations. Two of these showed an increase and two a decrease in the number of pencil lifts per maze. The fifth group, thalamotomy subjects, showed an increase in pencil lifts per maze.

There are two additional studies relating to the effect of brain damage on Q-score (Bennett, 1956; Sydow, 1953). Sydow reported that the Q-score of psychiatric patients diagnosed as being brain damaged was significantly elevated in comparison to patients diagnosed as psychotic. Bennett found no significant relationship between Q-score and Shipley-Hartford CQ, which is purported to be an indicator of cerebral dysfunction.

Q-score Conclusions

1. On the basis of the available evidence, there does not appear to be any relationship between frontal lobe damage and Q-score.
2. Other areas of the brain have been inadequately studied, and, while there is some indication that Q-score may be sensitive to certain types of brain damage, it is impossible at present to determine which types.

Work Skills and Employability

Test Age

Studies investigating the relationships between work skills and Porteus TA have used intellectually retarded subjects almost exclusively. A variety of criteria have been used in these investigations, which appear to relate in varying degrees to the ability to perform simple industrial tasks. The criteria include handwork ability, woodwork ability, ability to inspect gum boxes, and employment versus unemployment. The criteria are often based on the opinion of a single rater, whose degree of contamination is unclear. Further, most of the studies are not recent and their findings are based on out-dated versions of the Maze test. Some of the findings are reported as correlations; others as group differences in score. Despite all the flaws present in the investigations, the findings strongly suggest a positive relationship between TA and work skills that is somewhat greater than the relationship between work skills and intelligence tests such as the Binet.

In 17 separate investigations of the relationship using groups of retarded subjects, all indicate a positive relationship between Porteus TA and the criteria of work skills used. (Frost, 1965; Gambaro & Schell, 1966; Herd, 1923, 1930; Porteus, 1919, 1923, 1939b; Ross, 1921; Tobias & Gorelick, 1962). Porteus (1946) reported an unpublished study by Brundage in which a correlation of .50 was found between TA and ability to inspect gum boxes for a groups of 98 subjects, presumably of normal intelligence. Stotsky, Sacks and Daston (1956) published a study in which the relationship of the Maze Test to work skills is investigated,

but the version of the Maze test used in this study combined elements of TA and Q-score into a single composite score. This score probably bears little relationship to either TA or Q-score separately. This version of the test significantly differentiated a group of good and poor psychiatric hospital aides in a preliminary investigation, but failed to do so in a cross validation of the original findings.

The relationship between Porteus TA and work skills has been compared to that between Binet IQ and work skills among eight groups of subjects (Porteus, 1919, 1923, 1939b; Ross, 1921). In seven of these comparisons, work skills correlated higher with the Porteus than with the Binet. In the study by Ross, the correlations were equal. Tobias and Gorelick (1962) reported a higher correlation with work skills for the Porteus than for the WAIS and, for a separate group of subjects, found the Porteus to discriminate between individuals who were later employed versus unemployed, when WAIS IQ was held constant.

Test Age Conclusions

1. Although many of the studies in the area of work skills are old and poorly designed, they nevertheless strongly suggest a positive relationship between TA and simple work skills, at least at lower levels of intelligence.
2. The relationship between TA and work skills appears to be somewhat greater than that between intelligence tests and work skills.
3. The findings in this area are based almost solely on investigations with mentally retarded subjects and should not be generalized to non-retarded individuals without further research.

Q-score

Only two studies attempt to relate Maze Q-score to work skills or employability. Craft, et al., (1962) found the discharge Q-score of a group of 67 institutionalized delinquents to be unrelated to whether they were employed one year following discharge. Porteus (1945) administered the Maze test to summer employees of an Hawaiian pineapple cannery, mostly high school students. Subjects were judged by their foremen as having satisfactory or unsatisfactory work records. Unsatisfactory was defined as slow, indifferent, poor attitude, lazy, and/or some kind of discipline problem. The unsatisfactory workers had higher Q-scores than did the satisfactory workers.

Q-score Conclusion

Evidence for or against a relationship between Q-score and work skills or employability is inadequate for drawing conclusions.

Functional Psychiatric Patients

Test Age

A number of studies report TA results for psychiatric patients or emotionally disturbed children (Aaronson, 1963; Conners & Eisenberg, 1963; Conners, Eisenberg & Sharpe, 1964: Cornell & Lowden, 1923; Crawford, 1959; Crown, 1952; Docter, 1960b; Jones & Peters, 1952; Jones, 1949; Leiter, 1959; Michaels & Schilling, 1936a; Myatt, 1951; Peters & Jones, 1951; Porteus, 1957; Porteus & Barclay, 1957; Porteus & Kepner, 1944; Porteus & Peters, 1947; Sheer, 1956; Sheer, et al., 1952; Smith, 1960; Sydow, 1953; Weintraub, 1968; Zubin, et al., 1961 . The majority of the subjects in these

studies were hospitalized for one of the psychoses, most often schizophrenia, although subjects from a number of other diagnostic categories were investigated.

The most striking finding from the studies of schizophrenic subjects is the variability of TA across groups. The mean TAs reported for groups of schizophrenic subjects range from 4.25 (Sheer, et al., 1952) to 16.67 (Sheer, 1956). The reported standard deviations within groups appear to be relatively high. Standard deviations were reported for ten groups of schizophrenics and, of these, nine were higher than the level expected for normal subjects. The range was from 2.0 - 4.4 years. With the exception of the eight pseudoneurotic schizophrenics from Sheer's 1956 study who all obtained high TAs, there did not appear to be any relationship between the diagnostic group of the schizophrenic subjects and their mean TA. The overall mean TA of all 22 adult schizophrenic groups for which TA was reported is 12.0 for a total of 259 subjects. Seventeen of these groups obtained mean TAs below 14. This is probably an overestimate of schizophrenic Maze test performance because these subjects, for the most part, were selected for their ability to cooperate during testing. A random sample of schizophrenic subjects would, therefore, probably have a mean TA of less than 12 years.

No useful relationship between TA and IQ among schizophrenics is apparent in the studies reporting both scores. Out of 16 possible comparisons between TA and intelligence, TA was low with respect to IQ for 8 groups, roughly the same for 3 groups, and high with respect to IQ for 5 groups. The sampling bias mentioned above would tend to elevate scores, both on the Maze test and on

the intelligence test, but would not necessarily be expected to affect the relationship between the two tests.

The reported TA scores of 12 other psychotic groups, not included in the above discussion, were all within one year of 12 (Crawford, 1959; Crown, 1952; Michaels & Schilling, 1936a; Porteus, 1957; Porteus & Barclay, 1957; Porteus & Kepner, 1944; Porteus & Peters, 1947; Sydow, 1953). Ten of these groups were of mixed diagnosis and probably included a high percentage of schizophrenics. Michaels and Schilling (1936a) gave separate scores for two diagnostic groups. The mean TA for 23 manic depressives was 12.5, and for 15 paranoid states it was 12.9. Leiter reported a mean TQ of 58 (equivalent to a TA of 8.1) for a group of 92 psychiatric patients of whom 85% were diagnosed as psychotic. An unreported number of these subjects, however, were taking tranquilizing drugs which may have lowered the mean Porteus score for the group. Good (1957), however, found higher scores among two groups of mixed psychotics and neurotics. Her subjects scored 14.1 and 15.0 years of TA.

Conners and Eisenberg (1963), Conners, Eisenberg and Sharpe (1964), and Weintraub (1968), reported maze test results for emotionally disturbed children. In the studies by Conners, et al., 39 disturbed children, ages 7 to 15 years, with a mean age of 11.8 years obtained a maze TA of 10.4. These subjects, however, were below average in intelligence with a mean IQ of 85. In Weintraub's (1968) study, 66 emotionally disturbed 11 to 13 year old boys obtained a significantly lower Porteus TQ than a group of 26 normal boys. The difference between these groups remained significant with intelligence held constant.

The only TA for a group of neurotics found in the literature was 14.5 for 39 subjects in the study by Michaels and Schilling (1936). This finding suggests that neurotics may do as well as normals on the Maze test.

Cornell and Lowden (1923) assessed the Maze test's value as a diagnostic instrument using psychopathic, schizophrenic, retarded, and nonpsychotic psychiatric patients as subjects and Zubin, et al., (1961), investigated its value in prognosis using length of hospitalization of schizophrenic patients as a criterion. Both groups of investigators concluded that Porteus Maze TA was of little value in the areas studied. In a study by Myatt, (1951), however, Maze TA discriminated 18 patients showing low motivation during their rehabilitation, as rated by hospital staff, from 39 patients showing average motivation and 19 patients showing high motivation. The subjects with low motivation obtained a significantly lower mean TA than did the subjects in the other two groups.

Test Age Conclusions

1. Functional psychotic patients, in general, obtain lower Porteus Maze TAs than do normal subjects.
2. The Maze test's value as a diagnostic or prognostic instrument with nonorganic patients in the psychiatric setting is unclear and should be further investigated.
3. The need for further research into Maze test performance of emotionally disturbed children is indicated.

Q-Score

There are 13 studies reported in which Q-score data is given for nonpsychopathic and nonorganic psychiatric patients (Aaronson, 1963; Anthony, 1963a, 1963b; Craft, 1966; Crown, 1952; Docter, 1960b; Foulds, 1951, 1952; Jones & Peters, 1952; Minski & Desai, 1955; Peters & Jones, 1951; Porteus & Peters, 1947; Sydow, 1953; Weintraub, 1968). Both psychotic and neurotic groups have been investigated, but results are equivocal. Mean Q-scores are either high in comparison with a matched normal group or above the delinquent cut-off point of 29 for 9 out of 13 groups for which this information is available. The other 4 of these 13 groups do not appear to be above normal in Q-score. These studies include both neurotic and psychotic groups, but there is no apparent tendency for either neurotic or functional psychotic subjects to obtain the higher Q-scores.

Studies by Craft (1966), Peters and Jones (1951), and Jones and Peters (1952) indicate a tendency for Q-scores to decrease with improvement following psychotherapy. These studies were too limited to be more than suggestive and their findings need further confirmation.

Q-Score Conclusions

1. In the reported studies, no clear-cut relationship is apparent between Q-score and psychopathology among neurotic and psychotic psychiatric patients.
2. There is some indication of a tendency for Q-score to decrease with improvement following psychotherapy, suggesting a need for further investigation in this area.

Personality Tests

Test Age

Little information is available dealing with the relationship between the Maze test and tests of personality. What information is available is in such widely scattered areas that few generalizations are possible. Normal adult subjects rated well-adjusted on the Rorschach were found to score higher on the Mazes than poorly-adjusted subjects of equal intelligence in a study by Fisher and Sunukjian (1950). No difference in TA was found between normal children measured as high or low in defensiveness by Ruebush, et al., (1963) on their Defensiveness Scale for Children. Satter (1955) found a non-significant negative correlation between TA and aspiration on the Level of Aspiration Test ($r = -.20$) for 48 institutionalized retarded adult subjects. For this measure, the subject is asked to estimate how well he will perform a simple task before each of a number of repeated trials. His score is the sum of the differences between his estimated score and his actual score on the immediately preceding trial. Small (1954) found a significant correlation of -.22 between TA and neuroticism as measured by the Cornell Index among a group of 62 normal adults indicating lower TA scores among more neurotic subjects. Weintraub (1968) found a significant difference in TA between emotionally disturbed children rated as externalizers and those rated as internalizers on Achenbach's Symptom Checklist (Achenbach, 1966), the internalizers obtaining the higher scores. The difference remained significant when Maze scores were adjusted for differences in intelligence between the two groups.

Test Age Conclusion

It appears there is some low order relationship between TA and certain characteristics measured by personality tests that may be independent of intelligence and warrants further investigation.

Q-Score

Very little work has been done in the area of the relationship between Q-score and various aspects of personality as assessed by tests. Gibbens (1958) reported correlations between Q-score and MMPI scales for a group of 100 delinquents combined with 52 normal subjects. He found significant correlations between unweighted Q-score and depression (negative), schizophrenia (negative) and mania. Using weighted Q-score, only the negative correlation with depression was significant.

LaBarba (1965) found no significant correlations between Q-score and various Rorschach measures relating to color responses among 15 adult psychiatric patients. Ruebush, et al., (1963), found a significantly lower Q-score among 16 low defensive boys, as measured by the Defensiveness Scale for Children, than among 15 high defensive boys. Weintraub (1968) found a highly significant difference in Q-score between 26 boys classified as internalizers and 40 boys classified as externalizers on Achenbach's Symptom Checklist (Achenbach, 1966), the externalizers obtaining the higher Q-scores.

Craft, et al., (1962) reported that Q-score was related to a cluster of tests (e.g., MMPI Hy, Pd, Sc; Rorschach use of

color, etc.) measuring what they labeled "extraversion" within a group of 62 males with behavior disorders. Little meaning can be attached to this finding because the authors gave no information as to the degree of relationship or their method of assessing it.

Q-score Conclusion

It appears likely that there is a low order relationship between Q-score and certain personality test scores. While research is far from adequate in this area, a tentative generalization can be drawn that the correlations between Q-score and personality tests reflect, in general, the Q-score's sensitivity to individuals with a cluster of tendencies including impulsivity, extrapunitiveness and acting out.

Delinquency, Criminality, and Psychopathy

Test Age

In the 1930s, the idea was frequently expressed by Porteus in his writings that Maze performance, all else being equal, was poorer among subjects of poor social adjustment than among well-adjusted subjects. The idea originated with Porteus in his work with retarded children in the early 1920s and gained prominence following the publication of studies by Poull and Montgomery (1929) and Karpeles (1932). In both of these studies, highly significant differences in Maze TA and TQ were found between institutionalized children with delinquent characteristics and children considered to be socially adjusted. There were no group differences in intelligence as measured by the Binet in either

study. Further, in each study the adjusted subjects obtained a higher score on the Maze test than on the Binet, while the delinquent subjects obtained a higher score on the Binet than on the Maze test. It was therefore felt that the relationship between an individual's Binet and Maze test scores was an indication of his capacity for social adjustment and also of his propensity for delinquent behavior.

Since the 1930s, mention of this aspect of the Maze test has almost completely faded, apparently due to lack of interest rather than because of any controversy concerning its validity. With the introduction of Q-score in 1942, investigation of the relationship between TA and delinquency came to a halt. It is worthwhile to re-examine the interrelationships among Maze TA, intelligence and delinquency in an attempt to establish whether this aspect of the Maze test has any utility.

It was originally felt that the relationship believed to exist among TA, intelligence, and social adjustment would be of use in identifying delinquents. It is clear from the data available on delinquent groups that there is no consistent relationship between TA and intelligence within these groups (Erikson & Roberts, 1966; Erikson & Roberts, 1971; Grajales, 1945; Jarrett, 1926; Porteus, 1939b; Shakow & Millard, 1935; Wright, 1944). Erikson and Roberts (1966) found the mean TQs (using the old norms) of two delinquent groups to be over 20 points higher than their mean IQs as measured by the California Short Form Mental Maturity Test. Erikson and Roberts (1971) found the TQs of two groups of delinquents to be over 15 points higher than their Binet, WAIS or WISC IQs. Grajales (1945), Porteus (1939b),

Shakow and Millard (1935), and Wright (1944) all reported mean Maze scores that were within three points of the reported mean IQ test scores, and out of four comparisons TQ was higher than IQ in three. Jarrett (1926), on the other hand, stated that none of the 100 delinquents in his study had a Maze TA higher than his Binet MA. In general, these results certainly do not fit with the expectation that delinquents would rather consistently obtain higher scores on intelligence tests than on the Mazes.

Comparisons between Maze scores and intelligence test scores are apparently not helpful in differentiating psychopaths from other types of psychiatric patients. Crawford (1959) found TQs to be significantly lower than IQs on both sections of the WAIS for a group of psychopaths and also for a group of psychotics. Michaels and Schilling (1936a, 1936b) found psychopaths to obtain higher TA scores than Binet MA scores, although the difference in score amounted to less than a year. In general, the other patients in this study tended to score slightly higher on the Binet than on the Mazes.

Since there appears to be no consistent relationship between intelligence test scores and Maze scores among delinquents, the possibility that TA may be lower among delinquents than among normal subjects should be examined. Most investigators report Maze scores among delinquent groups that are lower than the norms for the general population, insofar as these are known (Grajales, 1945; Porteus, 1939b, 1954, 1955; Shakow & Millard, 1935; Wright, 1944). These studies involve a total of 10 delinquent groups. Inconsistency in the method of reporting maze test scores in these studies makes it impossible to estimate the mean for the ten groups.

There are a substantial number of studies reporting mean Maze scores for delinquent groups that are not below the level of the general population. For three groups of delinquent subjects, the TA is within one-half of 14 years, the TQ (old norms) is within four points of 100, or the TQ (new norms) is within four points of 114 (Porteus, 1955; Porteus & Gregor, 1963; Roberts & Erikson, 1968). Maze scores above the general population norms have been reported for some atypical groups of delinquents (Erikson & Roberts, 1966; Erikson & Roberts, 1971). In their 1966 study, Erikson and Roberts found that two groups of delinquent subjects who were institutional management problems and one group of delinquents who were well-behaved, scored above the general population norms in TQ (another group of well-behaved delinquents in this study obtained a mean TQ which did not differ from that of the general population). In one study (1971), these investigators found that both a group of institutionalized delinquents who had chosen to attend high school off the institution grounds (even though this choice might increase the length of their institutionalization) and a matched control from the same institution, scored above the level of the general population in TQ.

It appears from the above studies that Porteus TA, separately or in combination with intelligence test scores, cannot be used to discriminate between delinquents and nondelinquents. If TA is a measure of social adjustment, it must surely be along some dimension other than delinquency.

The evidence is similarly equivocal with respect to recidivism.

In two unpublished studies concerned with personality variables associated with recidivism in delinquent males following release from a training school, Roberts, Erikson, and Bacon found that TQ was lower for recidivists than non-recidivists in one study but failed to discriminate in a second follow-up study.

Some studies deal with the ability to delay gratification, an area that appears to be related to delinquency. In these studies, subjects are divided into groups according to whether they choose to receive a prize or reinforcer immediately, or choose to wait for one of greater value. In only one of the studies (Weintraub, 1968), is any effort made to objectively determine the relative values of the reinforcers to the subjects.

The findings with respect to TA are equivocal. Roberts and Erikson (1968), Kainer (1965), and Weintraub (1968) all found TQ to be higher in subjects who had more ability to delay gratification. Erikson and Roberts (1971) and David and Bochner (1967) found no differences between the delay groups and control groups with respect to TQ. The negative findings of David and Bochner relate to children in a culture that is different from those reported in the other studies (Australian aborigine children). The Erikson and Roberts (1971) study used an exceedingly complex definition of "delay of gratification." TA as a measure of the ability to delay gratification may be the best predictor in situations where the planning, judgment, and foresight are of the immediate, practical kind rather than long term.

Test Age Conclusions

1. There is no consistent relationship between intelligence test scores and Porteus TA among delinquents which differentiates them from nondelinquents.

2. Porteus TA is not consistently lower than the general population level among delinquents.
3. There is no way presently known to distinguish delinquents from nondelinquents using Porteus TA.
4. The relationship between TA and recidivism is unclear.
5. The relationship between TA and delay of gratification is unclear.

Q-Score

The Porteus Maze Q-score was originally developed to enable testers to discriminate delinquents and criminals from nondelinquents and noncriminals. The evidence accumulated since its introduction overwhelmingly indicates that Q-score does discriminate as it was intended.

Table 3 summarizes the results from all the studies reporting Q-score means for delinquents or criminals and for normal subjects

Insert Table 3 about here

age 14 or older. It should be noted that with the exception of the group of Filipino plantation laborers included in the 1945 study by Porteus, there is no overlap of any of the delinquent and nondelinquent means shown in this table. Those studies reporting the significance of the Q-score difference between their delinquent and nondelinquent groups (Docter & Winder, 1954; Fooks & Thomas, 1957; Porteus, 1942, 1945) all report significance beyond the .001 level, except Porteus (1945), who found a difference significant at the .05 level between Filipino plantation laborers and Filipino criminals. Clearly, delinquent and criminal groups do,

Table 3

Q-scores of Delinquent and Normal Groups

Investigator	Criminals or Delinquents N	M	SD	Normals N	M	SD
Craft, et al. 1962	67	(23)				
Docter & Winder 1954	60	[47	26	60	[25	20
		[(24.7)	(13.2)		[(13.7)	(9.9)
Fooks & Thomas 1957	50	40.0	18.4	50	22.2	14.8
Gibbens 1958	191	[35	22	52†	[14	11
		[(17.7)			[(6.5)	
Grajales 1945	60	56.18	29.18			
		(28.34)				
Porteus 1942	100	[49.35	25.19	100	[21.77	12.9
		[(25.2)			[(12.1)	
	100	[52.9	24.39	100	[24.7	14.5
		[(25.3)			[(13.0)	
	100	[56.7	26.21	100	[17.85	12.85
		[(27.5)			[(8.9)	
Porteus 1945				100	19.25	13.04
(The non-delinquent groups were quite highly selected for good behavior except the last, which was a group of Filipino plantation laborers, some of whom were illiterate)				100	23.17	13.74
				100	22.54	14.15
	50	48.44	21.56	31	13.66	10.4
	50	62.3	30.14	25	15.82	11.42
	100	58	32.4			
				25	14.3	7.7
				33	21.53	13.26
	26	60.04	31.95	32	42.38	22.12
Porteus 1954	120	42	22			

-Continued-

Table 3 (Continued)
Q-scores of Delinquent and Normal Groups

Investigator	Criminals or Delinquents N	M	SD	Normals N	M	SD
Porteus 1955				300	23.86	15.25
Porteus 1961	23	49		42	24	
	50	40		39	22	
Porteus & Gregor 1963	21	35.19	17.34			
Purcell 1956	18	42.94				
	14	58.57				
Sanderson 1945				50	21.04 (9.56)	14.68
Schalling & Rosén 1968	83	35.59 (18.60)				
Wright 1944	54	49.26 (23.69)	27.72			
Total	1270	47		1339	22	

(Numbers in parentheses refer to unweighted Q-score).

† No attempt was made to exclude delinquents from this sample. The authors estimate a probable delinquency rate of 10-15% in the population from which the sample was taken.

in fact, obtain significantly higher mean Q-scores than do normal subjects.

Another way to investigate the Q-score's ability to discriminate delinquents from nondelinquents is by determining the percentage of delinquents and nondelinquents correctly identified by the test. A male subject is generally predicted to be delinquent if he obtains a Q-score above 29 and as nondelinquent if he obtains a score of 29 or less. For female subjects the cutoff point is a score of 32. The findings of studies using this method are summarized in Table 4. On the basis of this table, it appears that the hit rate is slightly more than 70%. Very few measures in this area do as well.

Insert Table 4 here

There is some indication that a cutoff point of 29 may be somewhat high for British subjects. Craft, et al., (1962) and Gibbens (1958) are the only investigators who have done Q-score studies within the British population and their reported mean scores are somewhat lower than those found in comparable samples from the continental United States or Hawaii. Gibbens found a Q-score of 22 to be a better cutoff point in differentiating British delinquents from normals. Using 22 as a cutoff point, 66% of his delinquents and 74% of his normal subjects were correctly identified. These results are, at best, only suggestive because only two groups of British delinquents and one group of British normals have been investigated. Further research is needed to determine whether the British population differs from the American population in terms of Q-score.

Table 4

Identification of Delinquents and Normals by Q-score

Investigator	N	% Delinquents Correctly Identified	N	% Nondelinquents Correctly Identified
Craft, et al. 1962	67	63		
Docter & Winder 1954	60	70	60	70
Fooks & Thomas 1957	50	76 (74)	50	72 (72)
Gibbens 1958	191	49		
LaBarba 1965 (psychiatric patients with a history of acting out or trouble with the law)	9	89		
Porteus 1942	100	80	100	76
	100	80	100	75
	100	86		
Wright 1944	54	78		
Total	731	74	310	74

Note: Numbers in parentheses refer to unweighted Q-score (cutoff point of 16).

In addition to the evidence of the Q-score's power of discrimination provided from studies using the Vineland Revision of the Mazes, three studies have been done using the other series of Porteus Mazes (Fooks & Thomas, 1957; Porteus, 1955; Porteus & Gregor, 1963). Fooks and Thomas reported a mean Extension Series Q-score of 46.4 for a group of 50 delinquents and of 27.8 for 50 nondelinquents. However, 42% of the nondelinquents were misidentified, which is an unusually high false-positive rate. In 1955, Porteus compared the Q-score performance of 300 high school students to that of 100 delinquents and 40 criminals using the Extension Series. The mean Q-scores reported for these three groups were 23.84, 53.59, and 46.3 respectively. These scores are quite consistent with the results from studies using the Vineland Revision mazes. Porteus and Gregor (1963) reported the Extension Series and Supplement Series results for a group of 60 high school students who had experienced some previous trouble with the law (excluding minor traffic violations). This constitutes a rather borderline group with respect to delinquency. Their Extension and Supplement Series mean Q-scores were reported as 31.01 (SD=20.96) and 34.34 (SD=20.71), respectively.

Not only does Porteus' Q-score discriminate delinquents from nondelinquents, it also appears to discriminate degrees of delinquency within both delinquent and nondelinquent groups (Erikson & Roberts, 1966; Porteus, 1945; Purcell, 1956; Schalling & Rosen, 1968). Erikson and Roberts (1966) found a significantly higher Q-score among institutionalized delinquents who were management problems within the institution than among those who were well-behaved. These results were replicated on a second sample.

Roberts, Erikson, and Bacon in two unpublished studies examined the relationship of Q-scores to recidivism in delinquent males following their release from a training school. In the first study recidivism was defined in terms of reincarceration following release during a one to three year follow-up period. The mean Q-score for 10 recidivists was 50.6 (SD=23.50) while the mean for 10 non-recidivists was 27.4 (SD=12.33).

Because N was so small in that study, Roberts, Erikson, and Bacon collected data on an additional 68 consecutively released students from the training school. One year following their release, follow-up data were obtained. 32 boys were found to have had "no contact" with their court or probation offices, 21 boys had had "some contact" (i.e., warnings, arrests but no conviction, probation violations, etc.), and 15 boys had been recommitted to an institution and considered recidivists. The mean Q-scores of the three groups were 20.1 (SD=13.01), 32.6 (SD=13.12), and 41.0 (SD=23.3) respectively, and the group differences were highly significant.

These findings are particularly impressive because the discriminations are among subgroups of delinquents and because the total group of 68 Ss was already truncated. Boys at the training center who were not behaving and not making progress in reaching the goals established with them at admission were not recommended for release and were therefore not in the sample. The subject population therefore consisted of boys who were judged to be able to stay out of trouble after leaving the institution.

Porteus (1945) found significantly higher Q-scores among high school students who were considered discipline problems by their

teachers than among those who were not. This relationship was found in each of three different high schools. Male pineapple cannery workers who were unsatisfactory because of slowness, laziness, indifference, or difficult discipline obtained significantly higher Q-scores than did satisfactory workers. The same result was obtained using female pineapple cannery workers.

Purcell (1956) found higher Q-scores among a group of subjects with extensive patterns of delinquency and criminal acts than among a group who committed only minor infractions of the law. The subjects in this study were all army enlisted men who were nonorganic, nonpsychotic psychiatric referrals to a Veterans Administration mental hygiene clinic. The significance of the difference between the two groups of subjects was not reported.

Schalling and Rosen (1968) found significantly higher Q-scores among criminals with high ratings on psychopathy than among criminals with low ratings. The ratings were based on Cleckely's definition of psychopathy.

Some conflicting results with regard to psychopathy were reported by Foulds (1951, 1952). A group of psychopaths were not found to have higher combined lift pencil and crossed line scores than the neurotic subgroups in the 1951 study. In 1952, however, Foulds reported a higher lift pencil plus crossed line score among a group of psychopaths and hysterics than among a group of anxiety states, depressives, and obsessionals. Because Foulds did not administer the mazes in the standard manner, did not report the total Q-score, and did not specify his use of the term, "psychopath", his findings cannot be given much weight.

In general, the evidence supports the conclusion that within any group, those with more delinquent characteristics obtain higher Q-scores than those with fewer delinquent characteristics. Another interesting point brought out in the investigation of Q-score is that its ability to discriminate between groups does not appear to be improved by the Q-score weighting system proposed by Porteus. The data in Table 3 indicates that the mean weighted Q-score can be estimated accurately to within three points by doubling the unweighted Q-score, regardless of the characteristics of the group. This relationship is true for all 20 of the groups for which both weighted and unweighted Q-scores have been published (Docter & Winder, 1954; Gibbens, 1958; Grajales, 1945; Porteus, 1942; Sanderson, 1945; Schalling & Rosen, 1968; Wright, 1944). For two groups in a study by Erikson and Roberts (1966), the estimate, based on unpublished data (Erikson, 1965) is within six points of the actual weighted Q-scores, and for the other two groups, it is within three points of the weighted Q-score. There is some indication that weighted standard deviations can be estimated from unweighted standard deviations in the same manner. Docter and Winder (1954), and Fooks and Thomas (1957) report that the percentage of delinquents and nondelinquents correctly identified by Q-score is not affected by the Porteus weighting system. There are no published data available to determine whether the relationship between weighted and unweighted Q-score holds up for individual subjects.

An area of study somewhat related to the studies on delinquency and criminal behavior is the ability to delay gratification. In the research of delay of gratification surveyed here,

subjects are generally offered a prize or reinforcer immediately or one of greater value if subject waits for a period of time.

All the studies reporting Q-score data (Roberts & Erikson, 1968; Erikson & Roberts, 1971; Verrill, 1958; Weintraub, 1968) find the Q-score significantly lower for the delaying less-impulsive groups and higher for the more-impulsive, immediately-gratifying groups.

Q-Score Conclusions

1. Delinquents obtain higher Q-scores than nondelinquents.
2. Using weighted Q-score and a cutoff point of 29 for males and 32 for females, slightly better than 70% of subjects tested can be correctly identified as delinquent or non-delinquent.
3. Within a given group, those with more delinquent tendencies obtain higher Q-scores than those with less delinquent tendencies.
4. Recidivists among delinquent males appear to obtain higher Q-scores than nonrecidivists.
5. The system of weighting proposed by Porteus does not increase the ability of Q-score to discriminate between delinquents and nondelinquents.
6. Subjects who tend to delay gratification obtain lower Q-scores than do more impulsive subjects.

Social Adjustment

Test Age

Studies of social adjustment, using the Maze test, have been of two general types: Studies of the correlation between TA and

ratings of social adjustment (Dentler & Mackler, 1962; Porteus, 1919, 1920, 1923, 1924, 1933, 1939b; Roberts & Erikson, 1968) and studies of group differences in Maze TA between institutionalized individuals classified as being of good or poor institutional adjustment (Brill, 1937; Erikson & Roberts, 1966; Karpeles, 1932; Porteus, 1923; Poull & Montgomery, 1929; Roberts & Erikson, 1968). The majority of the studies of the former type were done by Porteus and were apparently based on data accumulated during his brief stay at the Vineland, New Jersey Training School for Mental Defectives just prior to and during the early 1920s. These studies show correlations fairly uniformly centering around .6 to .7. Porteus reports 14 correlations involving a total of over 400 subjects. The criterion of social adjustment employed by Porteus was always the rating of an individual staff member of the training school who worked closely with the subjects. Usually this rater was Mrs. Nash, the School Department Principal. For the most part, Porteus' Social Ratings Scale (Porteus, 1920) provided the guidelines for subject evaluation. The areas covered by this scale are impulse control, emotional control, capacity for decision making, ability to master new tasks, and freedom from gullibility.

For eleven of the groups used in the calculation of TA social adjustment correlations, correlations with the criterion were also reported for the Binet. While correlations with social adjustment for the two tests were roughly on the same level, the Maze test shows the stronger relationship in 9 of the 11 groups, and both tests yielded equal correlations in the remaining 2. The implication of these findings is that the Mazes predict social adjustment ratings slightly better than does the Binet.

Dentler & Mackler (1962) obtained a correlation of .64 between TA and a laboratory performance version of the Vineland Social Maturity Scale among a group of institutionalized retarded boys. This correlation became .58 with age held constant. Roberts and Erikson (1968) reported a correlation of .36 (N=50) between TQ and their measure of institutional adjustment which was based on ratings of the degree to which subjects constituted discipline problems. Although this correlation is statistically significant, it is lower than the correlations based on other criteria of social adjustment.

Studies of the second type, investigating group differences, do not show such good agreement. It may be that the lack of agreement in this area stems from differences in the definition of institutional adjustment employed in these studies; the definitions found in the literature tend to be rather vague and the differences between them unclear. It may also be the case that some of the definitions employed have little to do with social effectiveness. Three studies (Karpeles, 1932; Poull & Montgomery, 1929; Roberts & Erikson, 1968) found significant group differences indicating that subjects considered to be institutional management problems obtain lower TA scores than do well-adjusted subjects. Three other studies (Brill, 1937; Erikson & Roberts, 1966; Porteus, 1923) found group differences in the opposite direction for all eight groups investigated, but these differences failed to reach significance. The situation is further complicated by the fact that the subjects were delinquents in the two studies by Roberts and Erikson and retardates in the other four studies.

One other study appears in the literature concerning the Maze test's relationship to social effectiveness (Cooper, et al., 1967), but in this study differences in intelligence were probably more crucial than were differences in social adjustment. These investigators found a highly significant difference in TQ between Southern Negro adolescents institutionalized as mentally retarded, who were rated as behaviorally non-retarded, and those who were rated as behaviorally retarded; the non-retarded group obtaining the higher score.

Test Age Conclusions

1. Correlational studies of social adjustment indicate a highly significant degree of relationship between the Porteus and the usual criteria of social adjustment employed.
2. While it is possible that most of this covariance could be accounted for in terms of intelligence, findings indicate that the correlation with social adjustment is somewhat higher for the Porteus than for the Binet. Unfortunately, no investigator has reported a correlation between the Mazes and social adjustment with intelligence ruled out.
3. Studies investigating group differences in TA among subjects of varying degrees of institutional maladjustment produced no consistent findings.

Anxiety and Stress

Q-Score

A few Porteus investigators appear to interpret Q-score as indicating anxiety (e.g., Agnew & Agnew; 1963; Leibowitz, 1966).

There are, however, only four studies in the Maze literature that attempt to investigate the relationship between these two variables and the results are equivocal (Agnew & Agnew, 1963; Anthony, 1963a, 1963b; Sarason, et al., 1960; Sternberg, 1964). Anxiety is operationally defined in these studies in two different general ways. One type of anxiety studies is induced anxiety in which the test is administered in presumably anxiety-producing and nonanxiety-producing manners, and the performance of subjects under these two conditions is compared. The other type of anxiety is dispositional anxiety which is defined in terms of performance on tests reported to measure anxiety, and, in one study (Anthony, 1963a, 1963b) defined by institutionalization for neurosis. The performance of subjects rated high and low on these measures is compared.

Two studies (Agnew & Agnew, 1963; Anthony, 1963a, 1963b) found significant differences in performance indicating that induced anxiety elevates Q-score. Sternberg, on the other hand, reported lower Q-scores for both comparisons made in his study under conditions of induced anxiety than under neutral conditions, but the differences were not significant and the groups were small, four groups of six subjects each.

In the studies of dispositional anxiety only Anthony (1963a, 1963b) found a significantly higher Q-score among the more anxious subjects. Sternberg (1964) reported higher Q-score among the anxious subjects in one comparison and a lower Q-score in the other comparison; however, both findings were nonsignificant. Sarason, et al., (1960) reported three comparisons in which the more anxious subjects obtained the lower Q-scores. The statistical

significance of the findings was not reported.

The discrepant results are rather easily rationalized on the basis of differences in criteria used in these studies. The data, however, are inadequate to determine exactly how differences in criteria affected the results.

Q-Score Conclusions

1. The four studies relating Porteus Q-score to "anxiety" and "stress" are clearly inadequate to determine the effects of these variables upon Porteus performance. The operational definitions of anxiety used are sufficiently different to raise questions about the comparability of the independent variable. For this reason, the inconsistent findings are not surprising.
2. A high Q-score should not be interpreted to imply anxiety, particularly in view of the fact that the delinquent, impulsive, and psychopathic subjects clearly produce higher Q-scores than do traditionally more anxious individuals.

Chapter VI

Summary and Conclusions

Summary and Conclusions

The Porteus Maze Test has been a quietly persistent instrument. It has never really achieved the spectacular, dramatic notoriety that occasionally crops up among psychological tests, but the Maze test has had continuous, widespread use during the sixty years since it was first developed. There is something about solving a maze that captures the interest of all types of people. Immediate fascination with the test is found among such an unlikely collection of subjects, as African bushmen, kindergarten children, the Ainu of Japan, graduate students, residents of old age nursing homes, hospitalized schizophrenics, prison inmates, adolescents of all descriptions and many more. It is one of the few tests that can provide useful information about people, from stone age primitives to the most highly educated members of any culture, and from preschool children through the aged.

With over half a century of research behind the Maze Test, there is a broad range of information available. The test is a proven measure of intelligence, but not so much of verbal ability as of some kind of practical aptitude, and the test identifies individual differences in this aptitude among the mentally retarded as well as those of average intelligence.

Among the most striking achievements for the Maze test is its ability to identify delinquents. Without exception, highly significant differences have been found between criminal and noncriminal groups in every published study in this area. Hence, the Maze test has proved itself most convincingly, to provide crucial information about a core of problems currently crying for solution.

Another surprising accomplishment for the test arose from the extensive work with psychosurgery, which all but faded out in the early 1950's. The test was found to be acutely sensitive to frontal lobe damage in the brain in spite of the fact that most tests that were expected to reflect this type of damage were totally unaffected by the surgery. Today, the Maze test is one of the best psychological instruments available for localizing frontal lobe lesions.

Summary of Test Age Findings

Unlike most intellectual tests, the Porteus Maze test appears to be relatively free from the influence of demographic variables among the type of individuals likely to be seen for psychometric evaluation. Sex, old age, and social class have been found to be related to TA, but the findings do not rule out the possibility that this relationship is due to innate differences in ability among the groups tested as opposed to discrimination against some of the groups by the test. Findings reported in the literature strongly suggest that certain drugs can influence Maze performance, so that test results obtained from subjects taking such drugs should be interpreted with caution. Practice appears to increase TA and this fact should always be kept in mind by the test user. Subjects hospitalized for psychosis or institutionalized for retardation score below the general population mean in TA.

The Maze test has considerable face validity as a measure of foresight and planning ability. It appears to test the capacity of a subject to look ahead, attend to appropriate information, and make intelligent decisions based on that information. Exactly how the ability to exercise foresight and planning on the test relates to the ability to use foresight and planning in everyday life is a question that has not been adequately investigated. Whether people with low TAs run out of gas on the freeway, get lost in the big city, overdraw checking accounts or fail to come in out of the rain more often than do people with high TAs is really unknown.

The test is a demonstrated measure of intelligence. About one-fourth of the variance in TA can be accounted for in terms of intelligence test scores, but the Maze test apparently does not

measure that aspect of intelligence covered by vocabulary tests. Together, these findings imply that Maze test performance is rather strongly related to an aspect of intelligence that is independent of word knowledge; however, this implication has yet to be studied adequately.

Among current psychometric tests, the Maze test is particularly sensitive to frontal lobe injury in the brain. It is one of very few measures responsive to this relatively silent area and, as such, could be an invaluable tool in localizing lesions that might otherwise be difficult to detect. The test also appears to be sensitive to the functioning of certain portions of the diencephalon. The relationship between brain damage and TA has not been investigated among children, however, so extreme caution should be exercised in the interpretation of the TA of these subjects in the area of cerebral functioning.

Among institutionalized retarded subjects, the Maze test has been found to predict work skills and to predict social adjustment, defined in terms of ability to exercise emotional control and function adequately with relative freedom from supervision. Together, the findings of these studies with retardates imply that the Maze test should predict success in functioning outside an institution because the test appears to measure some of the more important abilities necessary for such success. The Maze test could play a useful role in decisions concerning the need for institutional care of retarded individuals. It is also possible that the test could be of use in similar decisions involving brain-injured subjects. The use of the Maze test in this area appears to be very promising and certainly deserves more thorough research.

The Porteus Maze test has been found to measure intelligence, certain types of brain damage and abilities relating to the need for institutional care and supervision. The test appears to make a unique contribution, to provide information beyond that supplied by other measures and yet, as currently administered, it has not been adequately standardized. Adult norms are not available. It is administered in a variety of idiosyncratic ways, some of which are known to affect the resultant scores. Before the full potential of the test can be realized, investigators must agree on a standard method of administration (Porteus' 1965 instructions would do nicely) for which norms must be derived. All three series, the Vineland, the Extension, and the Supplement should be included in the norms. When this is done, test results can be generated without all of the ambiguity and waste of information characteristic of so much of the Porteus Maze research literature.

Summary of Q-Score Findings

Porteus Maze Q-score is a reliably scored measure used to make rough classification of subjects along a personality dimension. Q-score appears to be almost exclusively related to a core of traits characterized by impulsivity, acting out, discipline problems and delinquency. The few test-retest and alternate form correlations available suggest it may be unwise to use Q-score in making fine discriminations. The reliability of this measure as it is actually used has never been tested but the accuracy with which it classifies subjects suggests its reliability must be fairly high. Q-score has not been demonstrated to be influenced by subject characteristics such as sex and ethnic background. Many possible influential

characteristics such as age and social class, however, have not been adequately investigated. Considering the current state of knowledge about Q-score, it should not be used clinically with subjects under 14 years of age. Although it is known that these young subjects obtain higher scores than do older subjects, the correlates of Q-score for the under 14 age group have not been investigated.

Every study testing the ability of Q-score to discriminate between comparable delinquent and normal groups has reported highly significant group differences. Q-score is found to make similar discriminations even within delinquent and "normal" groups. Thus, it has been demonstrated to be sensitive to the whole range of the trait dimension usually summarized by the term "impulsivity". It identifies the presence of these traits with sufficient accuracy to be considered a significant and valuable measure.

Q-score norms are inadequate, being based largely on the performance of adolescents. As is the case with TA, Q-score tends to be idiosyncratically scored to the extent that much potentially useful research information has been lost, and as with TA, standardized administration and scoring could correct this situation.

Suggestions for Future Research

This review suggests certain directions for future research with the Porteus Mazes. Current norms for both TA and Q-Score are critically needed. These norms should be based on deviation quotients as with the WAIS but, in addition, it may be necessary to establish separate norms for males and females. As part of

the development of norms, it would be useful to devise an empirically weighted system for the Q-score. A more accurate determination of the reliability of Q-score is also needed. The possibility of increasing the test's ceiling should also be considered.

Additional data are needed on the test performance of subcultures in the United States, such as blacks, American Indians, and those in various culturally deprived groups. This review would suggest that the test might prove useful in evaluating groups of individuals who tend to perform below average on the usual tests of intelligence (which do seem to predict success in middle class white public schools) but who otherwise may have the intellectual-cognitive capacities for a high level of adjustment and function.

Further studies of the effects of the various commonly administered psychopharmacologic agents on Maze test performance are likely to shed additional light on both the Maze test and the effects of the drugs. This might be particularly true if combined with studies of the effects of localized brain dysfunction upon Maze Test performance. The relationship of Maze performance to electroconvulsive therapy has not been adequately explored.

An investigation of the characteristics of nondelinquents who obtain high Q-scores and of delinquents who obtain low Q-scores might prove enlightening. More longitudinal study of Q-score's ability to predict future delinquency is called for.

Similar work should be done describing any consistent characteristics found among subjects at various levels of TA. Perhaps these subjects would differ in their ability to plan ahead, but until such comparisons are made, this is not a certainty.

It is also suggested that a longitudinal study of the relationship of Maze performance to the adjustment of retarded and brain-damaged individuals functioning outside of an institutional setting is needed. The ability of an immediate post trauma application of the Mazes to predict the future's adjustment of brain-damaged individuals should be further investigated. Many Maze test findings emmerged from studies employing subjects of very restricted age ranges; more work needs to be done to allow these findings to be generalized to all age groups, if such generalization is appropriate. Finally, research should be undertaken to determine the validity of clinicians' subjective judgments about subjects based on observations of their behavior during their Maze test performance.

Appendix I

Current Maze Test Application Instructions

The most recent Maze Test Application and Scoring Instructions published by Porteus (1965) are reproduced below for easy reference. These instructions are followed by a brief discussion intended to resolve a few ambiguous points contained in the instructions.

Complete 1965 Instructions (Porteus, 1965, pp. 244-255)

The Maze Test has currently three forms, known as the Original series or Vineland Revision (Porteus, 1919, 1933, 1950, 1959a), the Extension (Porteus, 1955), and the Supplement (Porteus, 1959b). It should be clearly understood that these are not alternative but supplementary series. The last two forms have been so devised and standardized as to give equivalent scores if applied in the order given above. The object in providing the three forms was to eliminate (or control) practice effects in such situations where the experimenter desired to repeat the tests. One example is research into the effects of drugs, whether stimulant, depressive or tranquilizing. Other kinds of therapy might call for reexamination to measure cumulative effects. Had there been practice-controlled forms available while psychosurgery was in fashion, the error of regarding postoperational effects as transient would have been avoided.

The only series which is standardized from three years

upwards is the Original. Because older subjects are much more often involved in therapeutic treatment and research, the Extension and the Supplement series begin with the VII Year design. Since the individuals to whom it is usually applied are older adolescents or adults, and have already worked through the Original series, they are familiar with the appearance of the test designs and know what is required of them. Therefore, leading up to the more difficult test items through experience with the simpler tasks set for young children is no longer necessary.

Summarized in the barest outline, the chronology of Maze Test publications is as follows:

The earliest test series, 1915.

A partially revised form with variations of scoring, The Vineland Revision, 1919.

A small book entitled Guide to the Porteus Maze Test, containing also Adult designs I and II, 1924.

A second book entitled The Maze Test and Mental Differences, 1933.

A monograph under the title of Qualitative Performance in the Maze Test, 1942.

A third book, The Porteus Maze Test and Intelligence, reproducing the tests with modification of instructions to allow of qualitative scoring, 1950 (reprinted in 1956).

A monograph, The Maze Test: Recent Advances, giving the standardization of the Extension Series, 1955.

A fourth book, The Maze Test and Clinical Psychology,

containing the Original and Extension series, as scored for test age and qualitatively, with new and formerly current tables of test quotients, 1959.

The Porteus Maze, Supplement Series, 1959, standardized against the Extension Series.

The present volume contains the three current series, Original, Extension and Supplement, setting forth variations in scoring and administration, with particular reference to application to primitive subjects in 1963. For the first time the three forms of the test are now printed together, and include the newest, most concise instructions for their application, the qualitative scoring and the tables of test quotients in current use.

That such a variety of references and cross references has been presented will not be surprising when it is remembered that the Porteus Maze Test has now been in use for fifty years and that during that time there have been some important modifications, not sufficient in extent, however, to change radically the appearance and character of the tests. In other words, the instrument, somewhat crude or tentative in 1914, has experienced further development and evolution by 1965. Changes that have taken place have come about through diversity in its environment of use or application in different fields, such as are described in this book. Without these developments and varied applications in some of psychology's most uncertain areas of functioning, the Porteus Maze Test would not have

achieved any maturity or longevity. It is the importance of the problems rather than the qualities of the measure that have ensured its survival and may favor its continuance. Where complete illumination is not possible, mankind uses what light there is.

But without uniformity of form and method, only confusion can result. It should be emphasized that the method and rules set forth herein are an essential part of the tests themselves. In other words, the rules make the game. Any methodological change, i.e., giving only one trial at each age level instead of the number prescribed, or adopting a different form of instructions given to the subject, could change the results in such a way that it would no longer be the Porteus Maze Test that is being applied. Strict adherence to the rules is therefore most decidely recommended, with only such minor departures from method as are approved in the test.

General Rules

Using the Original Series, children under six years of age begin the testing with the design for Year III. With older cases, unless mental deficiency is suspected, the testing begins at Year V. For timid or withdrawn subjects, with primitive individuals or special cases where the language used in instructions is a barrier to understanding the V and VI Year tests can be used for practice. Not until the latter maze design is performed correctly without help should the testing and the scoring of trials go on to the design for Year VII.

In the Extension and Supplement series there are no test designs below VII, and since the individual has already done the Original series, no practice or demonstration is allowed.

With a primitive adult or adolescent, or a frightened child of any ethnic group, it is permissible to have other persons familiar to him in the room, but only while demonstration and practice are in progress. When the Year VII test has been reached, arrangements should be made for observers to withdraw. Only in very rare instances should a parent or friend be allowed to remain and in that case he should sit in such a position as to be unable to see the subject's performance. Verbal comment should be strictly forbidden. Any help, verbal or otherwise, simply makes the subject more dependent.

Every effort should be made to establish a good relationship between the examiner and subject before the testing begins. When a strange adult, a teacher, or a person of a culture alien to the subject applies the tests, special efforts should be taken to establish this rapport.

The instructions for application and performance must be strictly adhered to. There are several errors commonly committed by examiners who are beginners in the use of the Maze. The first is leniency in applying the rule that no blind alley may be entered. Crossing by the width of an imaginary penciled line blocking off the

entrance constitutes an unsuccessful trial. It is quite common for a subject to make a slight deviation or turn opposite an opening as if intending to enter the blind alley, but if he corrects it in time, only the qualitative error of "changed direction" (CD) is recorded, and the test continues. If, however, the imaginary line is crossed, the test design is immediately removed and an unsuccessful trial is recorded.[1]

Another mistake in application is for the examiner to neglect to invert and repeat a test design if, after the subject has failed in the allotted number of trials in the test for the previous year, he succeeds in passing the next higher test. The rule is that this success - in case it was accidental - is not accepted. The test is reversed and the worse performance of the two presentations, ordinary or inverted, is recorded for scoring purposes. If, for example, a child fails Year X but passes the XI Year test in one trial, the XI Year must be inverted. If he

---

[1] The inexperienced examiner will sometimes excuse his leniency by saying that the subject allowed his pencil to slip. He has, however, no right to make any assumption of accidental error. Special cases, such as individuals with tremors or partial paralysis, may be accorded lenient treatment in regard to the above rule, but in their case only approximate test ages are expected or attainable.

required two trials, a half year is deducted from his score. Failure is recorded if there are two unsuccessful trials below XII, or four in tests above that level.

A third examining mistake is to point out the opening at the end of the VIII Year or a higher test. Frequently the individual will ask, "Where do I get out"? Except in the case of the V, VI and VII Year designs, the subject should be quietly told that he must find his own way out.[2]

Tracing the course with the pencil or finger in the air over the maze is strictly forbidden, as this amounts to a trial or an overt prehearsal. If the subject persists, he should be told to keep his hand by his side until he is ready to begin drawing through the maze.

Detailed Instructions

The examiner, preferably sitting opposite the subject, holds the top of the test design with the tips of his fingers so that it cannot be shifted in position. The surface of the table should be smooth and a moderately soft lead pencil of medium sharpness should be provided.[3]

---

[2] For very primitive subjects the VIII Year exit may be pointed out.

[3] A somewhat blunt pencil should be supplied by the examiner. If the point is too sharp, it may dig into the paper or it may make it easier for the subject to avoid touching the lines, an item that is scored qualitatively.

Year III. The examiner says:

"I want to see whether you can draw all around between these lines without crossing or touching them with the pencil. You draw just like this."

The examiner demonstrates a slow and careful drawing performance between the guide lines, beginning at the arrow and proceeding just around the first angle. He then places a new blank design on the table and hands the pencil to the child. Every encouragement should be given, and in some cases where necessary the examiner should put his hand on the child's so as to guide it around the first angle in the design. Three trials are allowed and the subject is allotted 3 year credit if on any trial not more than three lines are crossed.

Year IV. The examiner says:

"Do this the same way. Begin here (indicating starting arrow) and draw right around without crossing any of the printed lines."

Three trials are allowed, and after the second trial, the examiner may point out places where the lines were touched or crossed. Four year credit is allotted if on any trial there were not more than three line crossings.

Year V. The examiner says:

"This is what is called a maze and you must draw with your pencil like this." (Examiner takes the pencil and draws about 1.5 inches of the course from the starting arrow near the rat to around the first turn:) "These

lines are all supposed to be walls and this rat went in here (indicating arrow) to try and get some cheese." (Point to cheese at the end of the maze.) "Now I want you to draw a line showing me where the rat went to find the cheese. But you must be very careful not to cross any lines or to go into any place that is blocked at the other end. If you go into any blocked place, you cannot turn around and come out. You must start all over again with a new maze.

"One more thing you must remember - you can stop anywhere and look as long as you like, but try not to lift your pencil off the paper until you have drawn right to the end of the maze."

Two trials are allowed.

Year VI. The examiner says:

"This is another maze. Begin here and show me where the rat went to get the cheese. But do not cross any lines or go into any blocked places."

Two trials are allowed.

Note: The V Year and the VI Year tests are used for demonstration and forepractice in the testing of older illiterates or primitive subjects. For these individuals the reader is recommended to follow in general the procedures outlined in Dr. Gregor's field notes in Chapter 9. However, allowing observers in the testing area is not recommended except for the practice or demonstration designs. If subjects can be gathered together where the

examiner can demonstrate to the group the V and VI Year designs on a blackboard, that is an excellent way to smooth out misunderstandings and obtain rapport. In many examining situations this is impossible, but in either group or individual demonstrations, Gregor's introduction or some similar explanation is suitable, using pidgin English or dumb show. Miming can also be used with the deaf or in cases with extreme language difficulties.

Year VII. In the Original series for more sophisticated older children not under any suspicion of mental defect, this design is the starting point of the Maze Test application. It also marks the beginning of the Extension and Supplement forms, in which tests for years lower than VII are not provided.

For older individuals the "rat and cheese" introduction is no longer needed by way of explanation of the test. For cases who have already worked through the simpler mazes, there is, of course, no need to repeat the instructions. Otherwise, the VII Year test is introduced in the following terms.

The examiner says:

"I want you to suppose that this is a maze in the form of a street map. All the lines are stone walls. You can imagine, if you like, that you are walking or driving a car in here (examiner points to starting point marked S) and you have to find your way out here (examiner points to exit arrow). But you must be very careful not to bump into any

of the walls nor go into any blocked street, because if you do so you cannot turn around or back out. So if you go into a blind street, you must start all over again.

"This is not a speed test. You can stop anywhere as long as you like while you decide which way to go, but try not to lift the pencil off the paper until you are right outside the maze, and don't bump into any walls. Start as soon as you are ready."

Two trials are allowed.

Years VIII, IX, X. The examiner says:

"Begin here and find your way out." (Examiner points to the starting arrow, but not to the exit.)

Two trials are allowed in each test design.

Year XI. The examiner says:

"Begin here in the center and find your way out."

Two trials are allowed.

Years XII, XIV, and Adult I. The examiner repeats the instruction:

"Begin here in the center and find your way out."

Four trials are allowed in the XII, XIV, and Adult tests.

Note: As the tests get more difficult, it will sometimes happen, particularly in the Adult test, that the subject will pause and say, "There's no way out." If at the same time he lifts his pencil, this is scored as an unsuccessful trial, even though he may be on the right course, and the test blank is removed. If he does not lift his pencil after making this remark, the examiner should wait

a couple of seconds and then remove the design, substituting a new blank and proceeding with the testing until the allotted number of trials for that Year have been given.

SCORING PROCEDURES

Continue testing until all the designs of a series have been successfully worked through within the allowable number of trials. At any point where a subject draws through an imaginary line across the entrance to a blind street or alley, the design is removed and an "unsuccessful trial," not a failure, is recorded. Failure is recorded only if this takes place after the number of trials allowed in the rules has occurred - two in each test design up to and including Year XI, four in the XII, XIV, and Adult. Testing and scoring normally cease after three failures anywhere in the series have been recorded, or two successive failures in Year XI or above.

Occasionally, especially with delinquents, when a complete qualitative record is desired, the rule of discontinuing the testing after three tests have been failed may be relaxed. For example, a subject who failed at IX, XI, and XIV Year levels may be given a chance to do the Adult test. This extra performance is discarded in the test age scoring but can be used in the qualitative score. Delinquents are thus given the chance to augment their chance to accumulate the penalties accruing by reason of Q-score errors.

Crossing or touching lines in the process of drawing is a qualitative, not a test age error, unless a blind

alley is entered. Cutting across from one alley to the next to avoid drawing around to reach an opening is scored as a test age error and is at once recorded as an unsuccessful trial.

The general scoring rule is to take as the ceiling of the test the level of the highest test passed in the allowed number of trials, then deduct a half year for every unsuccessful trial. Since four trials are allowed in the XII, XIV, and Adult I test, a subject can lose or gain, according to the number of trials required on each, as much as two years in test age score on each of these higher tests. Below Year XII, since there are only two trials allowed in each test, each failure (i.e., two unsuccessful trials) incurs a loss of one year.

There are no separate designs for thirteen and fifteen years, but two-year credit can be earned if Years XII and XIV are passed on the first trial. If the subject has passed Year XII, but no higher test, his ceiling score from which deductions are made is 13 years. If he passes Year XIV, but fails the Adult, his ceiling is 15 years.

In the case of an inverted test (after failure in the test immediately preceding), the worse record of the two applications is used for scoring.

Among erroneous procedures occasionally followed by inexperienced examiners is that of allowing a subject more than the allotted trials. It is forbidden on the grounds that the extra practice may affect subsequent performance.

Scoring Example 1

Year	Trials	Deductions	Year	Trials	Deductions
VII	1		XI	1	
VIII	2	1/2 year	XII	2	1/2 year
IX	1		XIV	4	1-1/2 years
X	2	1/2 year	Adult	2	1/2 year

Ceiling score = 17 years less 3-1/2 years (7 unsuccessful trials); Test Age = 13-1/2.

Scoring Example 2

Year	Trials	Deductions	Year	Trials	Deductions
VII	1		XI Inverted	2	1/2 year
VIII	2	1/2 year	XII	Failed	2 years
IX	1		XIV	1	
X	Failed	1 year	XIV Inverted	4	1-1/2 years
XI	1		Adult	3	1 year

Ceiling score = 17 years less 6-1/2 years (13 unsuccessful trials); Test Age = 10-1/2. When a test is inverted, the worse performance is scored.

Scoring Example 3

Year	Trials	Deductions	Year	Trials	Deductions
VII	1		XI	Failed	
VIII	Failel	1 year	XII	Failed	
IX	2	1/2 year	XIV	3	No credit
IX Inverted	1		Adult	3	No credit
X	1				

Highest test passed = 10 years less 1-1/2 years; Test Age= 8-1/2 years. Test continued for qualitative scoring purpose , but no credit given for Adult and Year XIV successes, because of two successive failures in Years XI and XII.

Scoring Example 4

Year	Trials	Deductions	Year	Trials	Deductions
VII	1		XI	1	
VIII	1		XII	Failed	2 years
IX	Failed	1 year	XIV	2	
X	2	1/2 year	XIV Inverted	2	1/2 year
X Inverted	1		Adult	Failed	No credit

Ceiling score = 15 years less 4 years (8 unsuccessful trials); Test Age = 11 years.

Scoring Example 5

Year	Trials	Deductions	Year	Trials	Deductions
VII	1		XI	2	1/2 year
VIII	2	1/2 year	XII	3	1 year
IX	2	1/2 year	XIV	Failed	No credit
X	1		Adult	Failed	No credit

Ceiling score = 13 years less 2-1/2 years (5 unsuccessful trials); Test Age = 10-1/2 years.

Note: To convert test ages into IQ's, 14 years should be used as the divisor for cases at that age or above.

QUALITATIVE SCORING

The qualitative as distinct from the mental age scoring is concerned with errors in drawing or execution rather than in planning. In other words, it has to do with the subject's ability to follow instructions with regard to crossing lines and lifting the pencil, and also takes into account the general neatness of the performance and whether quantitative errors (blind street entrances) took place near the beginning or end of the test design. Qualitative scoring usually begins with Year VII.

The score is intended to reveal any haphazard, impulsive, or over-confident habits of action, or a tendency to become so absorbed in the task of finding the way out of the maze as to neglect other directions for executive performance. The weighting of errors is based on the comparative performance of delinquents, for whom the qualitative score was arranged.

No criticism of the performance is given by the examiner but every error in execution must be rigorously scored. Leniency may change the picture considerably. The higher the Q-score, the worse the record.

DETAILS OF SCORING

The following errors are marked on the face sheet, weighted as indicated and are then totalled for the qualitative score:

1. Any blind alley entrance (test age error) occurring in the first third (FT) of the design, i.e., at or

before the arrow marked on each sample of a qualitatively scored maze. Weighting 2.

2. Any blind alley entrance occurring in the last third (LT) of the design, i.e., at or after the arrow marked on the sample mazes. Weighting 1.
3. Cut corner (CC). The error must occur while actually turning a corner. Weighting 1. (A line crossed near a corner but not related to turning is scored as a crossed line.)
4. Wherever the pencil mark touches a printed line in a maze, other than in turning a corner, it is scored as a crossed line (CL). Touching a line when going through an opening is similarly scored. Weighting 2.
5. Lifting the pencil (LP) except at the exit is a qualitative error. The warning against lifting the pencil is given in the initial instructions for Test V (usually children 12 years of age or under) and in the instructions for Test VII with older cases. The warning should be repeated again after five lift pencils in a single test, or after a total of ten liftings in any combination of tests. This is the only allowable interruption on the part of the examiner while testing is in progress. The reason is to prevent the too early accumulation of error points up to the prescribed limits. Weighting 3.

6. "Wavy lines" (WL) are recorded against the subject if his general performance is as irregular in appearance or worse than the performance given as a scoring sample of the test in question. Weighting 2.
7. "Changed direction" (CD), formerly wrong direction (WD), is the notation made when it is evident that the subject started to go into a blocked path but avoided the error before crossing an imaginary line across a blind alley opening. Weighting 1.
8. Any qualitative error occurring in Tests VI and VII obtains an added penalty of one point for each such error. A place in the face sheet provides space for recording. Weighting 1.

The Q-score is the sum of the weighted scores in the above list, but there are two important restrictions. The maximum weighted error score recorded against any individual in any single type of error is 48. This means that only 24 crossed lines or 16 lift pencils are counted. It was found that in some cases, particularly delinquents, so many errors occurred that the score lost meaning for comparative purposes. There are no real distinctions between an individual scoring 100 and one scoring 150. For that reason the maximum recorded Q-score is 100.

Obviously, the subject should not be aware that his performance is being scored. However, the number of lift pencils in any single design should be counted and the

notation LP followed by the number should be marked on each test design where they occurred. In some cases the individual lifts his pencil but puts it down again at the same place so that the full number of LPs might not be apparent by later inspection. The "lift pencils" can be summated at the end of the testing.

CHANGES IN FORMER PROCEDURES

In test age scoring, the rule regarding successive failures in two year tests now applies above Year IX instead of Year VIII. This gives some individuals a better chance to adjust to the Maze situation.

The table for reckoning test quotients in the Maze allows cases to obtain a maximum TQ of 135 at each age level. In my opinion, TQ's below 30 and above 135 are of little comparative significance. The practice of including months of chronological age in divisors to obtain ratios between mental ages and chronological ages also seems unnecessary, as it confers an air of exactitude that is not justifiable. In the tables of TQ's appearing in the present volume, chronological ages are given in half years up to and including ten years,in 12-month intervals thereafter.

MAZE TEST EXTENSION

The Maze Test Extension series consists of eight test designs published in 1955 so as to give scores equivalent to those of the Original or standard series. This extension was devised in order to control practice effects whenever

it is necessary to repeat the Maze after an initial application. As an example of this use, we may instance studies of patients before and after some special therapeutic measures, such as psychosurgery or administration of tranquilizing drugs, where a practice-free form of the Maze is required.

EXTENSION SUPPLEMENT

A second practice-free series called the Extension Supplement has also now been standardized and published. Testing with the Extension and Supplement series begins at Year VII, and the instructions for application, scoring, and TQ reckoning are those set forth in the present volume.

Discussion of 1965 Instructions

Although nearly every possible test situation is covered by the current instructions, there are a few ambiguous points to be resolved, and there are some points that should have greater emphasis. These points are discussed in the order in which they occur in Porteus' instructions.

Porteus mentions in passing, early in the introductory portion of his instructions, that the two new practice-free series of mazes will give scores equivalent to the original series only if they are applied in the proper order. It should be stressed that valid Maze test scores can be obtained from the three series only if the Vineland Revision is administered first, the Extension Series second and the Supplement Series third. Any other order of presentation invalidates the test.

Throughout the entire history of the Maze test, Porteus

has never mentioned what the examiner should do if subject fails the first maze presented in the Vineland Revision, Year V for children and Year VII for adults. He does imply, however, that in cases where mental deficiency is suspected, testing should begin with the Year III maze, and it is our considered opinion that it would behoove the examiner to suspect mental deficiency, and begin with the Year III maze, following failure on the first maze applied. If an adult, for example, failed to successfully trace through the Year VII maze of the Vineland Revision in the allotted two trials, the rule would be to score a failure for the Year VII maze and administer the mazes for years III through VI before going on to the Year VIII maze. No such leniency would be allowed on the Extension and Supplement Series, however. The subject presumably knows how to do mazes before these two series are applied.

Porteus has been inconsistent, over the years, in his recommendations concerning how many mazes should be used to obtain Q-score. In his own Q-score validating studies and in most of the validating studies done by other investigators, the entire series of mazes was applied regardless of failure and Q-score was based on the performance on all of the mazes. In the early studies, prior to 1955, the Adult maze was not to be included in the Q-score, and in later studies the Adult maze was to be included. Q-score appeared to be equally valid by either method.

The current scoring instructions leave the decision of which mazes to include in the Q-score up to the individual examiner. Testing can be discontinued when the TA ceiling is reached and Q-score must then be based on the mazes administered

up to that point. If the examiner wishes, he can administer the remaining mazes and calculate Q-score for the performance on the entire series.

Although it is true that Q-score seems "magically" to work no matter how it is derived, it does seem like a good idea to establish a standard procedure and stick to it. Because most, if not all, of the studies done to validate Q-score used the performance on all the mazes, it would be best to routinely apply the entire series of mazes and calculate Q-score on the basis of the performance on all the mazes.

In the section on qualitative scoring, Porteus clearly states that a high Q-score is the result of an inferior performance. People are apparently so thoroughly conditioned into thinking that all high scores indicate good performance that Q-score still causes confusion. It should therefore be emphasized that Q-score goes in the opposite direction to our expectations.

Porteus clearly implies, but never directly states, that Q-score is to be based on every maze used by the subject whether the maze was successfully completed or not. If subject lifts his pencil four times on a maze and then enters a blind alley, the four pencil lifts are to be counted against him in the Q-score.

There is one exception to the rule that Q-score errors are to be counted for every maze. The "wavy lines" error is scored only once for each maze year, and is based on the general appearance of all the mazes for a given year. This rule was clearly stated in the 1942 scoring instructions. In subsequent instructions, the wording has been ambiguous, but there is no indication that Porteus ever intended to change the scoring of wavy lines,

and it is reasonable to assume that the 1942 scoring instructions still apply.

The TQ table provided by Porteus has not been reproduced here. It is our feeling that the test is badly in need of new norms, and use of the current TQ table is virtually unjustifiable.

The 1965 Maze application and scoring instructions provided by Porteus, with the few remaining areas of ambiguity resolved, constitute an adequate basis for uniform Maze test administration. If such a set of instructions could be universally adopted and employed in future Maze test administration, much of the confusion resulting from the variability in past Maze test administration would be resolved.

Bibliography

Bibliography

Aaronson, B. S. The Porteus Mazes and Bender-Gestalt recall. Journal of Clinical Psychology, 1957, 13, 186-187.

Aaronson, B. S. The comparative sensitivity of the qualitative and quantitative Porteus Maze scores to drug effects. Journal of Clinical Psychology, 1963, 19, 184-185.

Achenbach, R. M. The classification of children's psychiatric symptoms: A factor analytic study. Psychological Monographs, 1966, 80, (7, Whole No. 615).

Agnew, N. & Agnew, M. Drive level effects on tasks of narrow and broad attention. Quarterly Journal of Experimental Psychology, 1963, 15, 58-62.

Anthony, A. A. The relationship between anxiety and Qualitative Porteus Maze performance. (Doctoral dissertation, Columbia University) Ann Arbor, Michigan: University Microfilms, 1963. No. 63-7413. (a)

Anthony, A. A. The relationship between neuroticism, stress, and Qualitative Porteus Maze performance. Journal of Consulting Psychology, 1963, 27, 513-519. (b)

Arnold, G. F. A technique for measuring the mental ability of the cerebral palsied. Psychological Service Center Journal, 1951, 3, 171-178.

Aronson, H. & Klee, G. D. Effect of Lysergic Acid Diethylamide (LSD-25) on impulse control. Journal of Nervous and Mental Disease, 1960, 131, 536-539.

Arthur, G. A point scale of performance tests. New York: The Commonwealth Fund, Division of Publications, 1930.

Arthur, G. A point scale of performance tests (data that form the basis for the standardization of a point scale of performance tests). St. Paul: Typewritten, 1933. (a)

Arthur, G. A point scale of performance tests. Vol. II. The process of standardization. New York: Commonwealth Fund, 1933. (b)

Arthur, G. A point scale of performance tests. Vol. I. Clinical manual. (2nd ed.) New York: The Commonwealth Fund, 1943.

Arthur, G. A point scale of performance tests. Revised Form II. Manual for administering and scoring the tests. New York: The Psychological Corporation, 1947.

Ball, R. J. An objective measure of emotional instability. Journal of Applied Psychology, 1929, 13, 226-256.

Barry, J. R., Everstine, L., & Kleman, J. P. An abbreviated Qualitative score for the Porteus Mazes. Journal of Clinical Psychology, 1961, 17, 291.

Beach, H. D. & Johnston, C. A delayed reaction test for humans. Bulletin of the Maritime Psychological Association, 1960, 9, 11-21.

Bennett, H. J. The Shipley-Hartford scale and the Porteus Maze test as measures of functioning intelligence. Journal of Clinical Psychology, 1956, 12, 190-191.

Berry, R. A. & Porteus, S. D. Intelligence and social valuation. Vineland, N. J.: Training School Research Publication No. 20, 1920.

Blackwood, B. A study of mental testing in relation to anthropology. Mental Measurement Monographs, 1927, 4.

Briggs, P. F. The validity of the Porteus Maze test completed with the non-dominant hand. Journal of Clinical Psychology, 1963, 19, 427-429.

Brill, M. The validity of the Porteus Maze test. Journal of Educational Psychology, 1937, 28, 481-493.

Bronner, A. F., Healy, W., Lowe, G. M., & Shimberg, M. A manual of individual mental tests and testing. Boston: Little, Brown & Company, 1927, 140-141.

Brown, I. A., French, L. A., Ogle, W. S., and Jahnson, S. Temporal lobe epilepsy; its clinical manifestations and surgical treatment. Medicine, 1956, 35, 425-459.

Burt, C. The Porteus Maze test. In Burt, C. Mental and scholastic tests. London: The London County Council, 1921, 242-256.

Burt, C. Handbook of tests for use in schools. (2nd ed.) London: Staples Press, 1948, 98-110.

Conners, C. K. & Eisenberg, L. The effects of methylphenidate on symptomatology and learning in disturbed children. American Journal of Psychiatry, 1963, 120, 458-464.

Conners, C. K., Eisenberg, L., & Sharpe, L. The effects of methylphenidate (Ritalin) on paired-associate learning and Porteus Maze performance in emotionally disturbed children. Journal of Consulting Psychology, 1964, 28, 14-22.

Cooper, G. D., York, M. W., Daston, P. G., & Adams, H. B. The Porteus test and various measures of intelligence with southern Negro adolescents. American Journal of Mental Deficiency, 1967, 71, 787-792.

Cornell, E. L. & Lowden, G. L. A comparison of the Stanford and Porteus tests in several types of social inadequacy. Journal of Abnormal and Social Psychology, 1923, 18, 33-42.

Covault, R. E. The influence of test anxiety and task complexity on the learning of mazes. Dissertation Abstracts, 1955, 15, 1438.

Craft, M. Ten studies into psychopathic personality. Bristol, England: John Wright & Sons Ltd., 1965.

Craft, M. Methods of Care: II. The Balderton psychopathic unit. Chapter 7 in Craft, M. (Ed.), Psychopathic disorders and their assessment. New York: Pergamon Press, 1966, 116-134.

Craft, M., Fabisch, W., Stephenson, G., Burnand, G. & Kerridge, D. 100 admissions to a psychopathic unit. Journal of Mental Science, 1962, 108, 564-583.

Crawford, P. L. The relative sensitivity of the LAIS, WAIS, and PM in differentiating between psychopathic and psychotic patients in a mental hospital. Psychological Service Center Journal, 1959, 11, 93-97.

Crown, S. An experimental study of psychological changes following prefrontal lobotomy. The Journal of General Psychology, 1952, 47, 3-41.

Cunningham, K. S. Binet and Porteus tests compared. Examination of one hundred school children. Journal of Educational Psychology, 1916, 7, 552-556.

David, K. H. Effect of verbal reinforcement on Porteus Maze scores among Australian aborigine children. Perceptual and Motor Skills, 1967, 24 (3, pt. 1), 986.

David, K. H. & Bochner, S. Immediate vs. delayed reward among Arnhem-land aborigines. The Journal of Social Psychology, 1967, 73, 157-159. (a)

David K. H. & Bochner, S. Teacher ratings of IQ and Porteus Maze scores of Pitjandjara children. Perceptual and Motor Skills, 1967, 25, 639-640. (b)

Davies, A. D. & Davies, M. G. The difficulty and graded scoring of Elithorn's Perceptual Maze Test. British Journal of Psychology, 1965, 56 (2-3), 295-302.

Davies, M. G. & Davies, A. D. Some analytical properties of Elithorn's Perceptual Maze. Journal of Mathematical Psychology, 1965, 2, 371-380.

Dentler, R. A. & Mackler, B. The Porteus Maze test as a predictor of functioning abilities of retarded children. Journal of Consulting Psychology, 1962, 26, 50-55.

DeSanctis, S. Visual apprehension in the maze behavior of normal and feebleminded children. Journal of Genetic Psychology, 1931, 39, 463-468.

Docter, R. F. Interrelationships among Porteus Maze test qualitative errors. Journal of Clinical Psychology, 1960, 16, 336-338. (a)

Docter, R. F. Test-retest performance of schizophrenics on two forms of the Porteus Mazes. Journal of Clinical Psychology, 1960, 16, 185-187. (b)

Docter, R. F. & Winder, C. L. Delinquent vs. non-delinquent performance on the Porteus Qualitative Maze test. Journal of Consulting Psychology, 1954, 18, 71-73.

Doll, E. A., Phelps, W. M. & Melcher, R. T. Mental deficiency due to birth injuries. New York: MacMillan Co., 1932.

DuBois, C. Results of the Porteus Maze tests. Chapter 19 in DuBois, C., The people of Alor. Minneapolis: University of Minnesota Press, 1944, Pp. 552-555.

Earle, F. M. & Milner, M. The use of performance tests of intelligence in vocational guidance. London: H. M. Stationery Office, 1929.

Ebert, E. & Simmons, K. The Brush Foundation study of child growth and development. I. Psychometric tests. Monographs of the Society for Research in Child Development, 1943, 8 (2).

Elithorn, A. A preliminary report on a perceptual maze test sensitive to brain damage. Journal of Neurology, Neurosurgery, and Psychiatry, 1955, 18, 287-292.

Elithorn, A., Jones, D., Kerr, M., & Lee, D. The effects of the variation of two physical parameters on empirical difficulty in a perceptual maze test. British Journal of Psychology, 1964, 55, 31-37.

Elithorn, A., Kerr, M., & Jones, D. A binary perceptual maze. American Journal of Psychology, 1963, 76, 506-508.

Elithorn, A., Kerr, M., & Mott, J. A group version of a perceptual maze test. British Journal of Psychology, 1960, 51, 19-26.

Erikson, R. V. An investigation of two delinquent groups differing in institutional adjustment. Unpublished master's thesis, New Mexico Highlands University, 1965.

Erikson, R. V. & Roberts, A. H. A comparison of two groups of institutionalized delinquents on Porteus Maze test performance. Journal of Consulting Psychology, 1966, 30, 567.

Erikson, R. V. & Roberts, A. H. Some ego functions associated with delay of gratification in male delinquents. Journal of Consulting Psychology, 1971, in press.

Fick, M. L. The educability of the South African native. South African Council for Educational and Social Research, Research Series Number 8, 1939.

Fisher, S. & Sunukjian, H. Intellectual disparities in a normal group and their relationship to emotional disturbance. Journal of Clinical Psychology, 1950, 6, 288-290.

Fooks, G. & Thomas, R. R. Differential Qualitative performance of delinquents on the Porteus Maze. Journal of Consulting Psychology, 1957, 21, 351-353.

Foulds, G. A. Temperamental differences in Maze performance: Part I. Characteristic differences among psychoneurotics. British Journal of Psychology, 1951, 42, 209-217.

Foulds, G. A. Temperamental differences in Maze performance: Part II. The effect of distraction and of electroconvulsive therapy on psychomotor retardation. British Journal of Psychology, 1952, 43, 33-41.

Foulds, G. A. & Caine, T. M. Psychoneurotic symptom clusters, trait clusters, and psychological tests. Journal of Mental Science, 1958, 104, 722-731.

Foulds, G. A. & Caine, T. M. Symptom clusters and personality types among psychoneurotic men compared with women. Journal of Mental Science, 1959, 105, 469-475. (a)

Foulds, G. A. & Caine, T. M. The assessment of some symptoms and signs of depression in women. Journal of Mental Science, 1959, 105, 783-787. (b)

Foulds, G. A., Owen, A. Speed and accuracy on mazes in relation to diagnosis and personality. British Journal of Social & Clinical Psychology, 1964, 3, 34-35.

Frost, B. P. The Porteus Maze test and manual proficiency. Alberta Journal of Educational Research, 1965, 11, 17-20.

Fry, H. K. & Pulleine, R. H. The mentality of the Australian aborigine. Australian Journal of Experimental Biology and Medical Science, 1931, 8, 153-167.

Gambaro, S. & Schell, R. E. Prediction of the employability of students in a special education work-training program using the Porteus Maze test and a rating scale of personal effectiveness. Educational and Psychological Measurement, 1966, 26, 1021-1029.

Gaw, F. Performance tests of intelligence. Medical Research Council Industrial Fatigue Research Board Report #31. London: H. M. Stationery Office, 1925. (a).

Gaw, F. A study of performance tests. British Journal of Psychology, 1925, 15, 374-392. (b)

Gibbens, T. C. N. The Porteus Maze test and delinquency. British Journal of Educational Psychology, 1958, 28, 209-216.

Gibson, H. B. The Spiral Maze; a psychomotor test with implications for the study of delinquency. British Journal of Psychology, 1955, 64, 219-225.

Good, P. E. K. A psychological study of the effects of regressive electroshock therapy. Doctoral dissertation, University of Minnesota, 1957.

Grajales, M. L. Porteus' Qualitative Maze test as a measure of delinquency. Unpublished master's thesis, Fordham University, New York, 1945.

Gregor, A. J. & McPherson, D. A. The correlation of the Porteus Maze and the Gestalt continuation as personnel selection tests of peripheral peoples. Journal of Psychology, 1963, 56, 137-142.

Gruetzow, H. & Brozek, J. Intellectual functions and restricted intakes of B-complex vitamins. American Journal of Psychology, 1946, 59, 358-381.

Grygier, P. & Waters, M. B. Chlorpromazine used with an intensive occupational therapy program. A. M. A. Archives of Neurology and Psychiatry. 1958, 79, 697-705.

Havighurst, R. J. & Hilkevitch, R. R. The intelligence of Indian children as measured by a performance scale. Journal of Abnormal and Social Psychology, 1944, 39, 419-433.

Havighurst, R. J. & Janke, L. L. Relations between ability and social status in a midwestern community. I: Ten-year-old children. Journal of Educational Psychology, 1944, 35, 357-368.

Helper, M. M., Wilcott, R. C. & Garfield, S. L. Effects of chlorpromazine on learning and related processes in emotionally disturbed children. Journal of Consulting Psychology, 1963, 27, 1-9.

Herd, H. The Porteus Maze mental test. The Medical Officer, 1923, 30, 267-268.

Herd, H. The diagnosis of mental deficiency. London: Hodder & Stoughton, 1930.

Heston, J. C. & Cannell, C. F. A note on the relation between age and performance of adult subjects on four familiar psychometric tests. Journal of Applied Psychology, 1941, 25, 415-419.

Husband, R. W. Certain age effects on maze performance. Journal of Genetic Psychology, 1930, 37, 325-328.

Jarrett, R. F. Some observations on social capacity. Application of the Porteus Maze tests to 100 Borstal lads. The Lancet, 1926, 211, 1059-1060

Jensen, M. B. Alternation learning and mental pathology: Test procedures and findings. The Journal of Psychology, 1960, 50, 211-225.

Jensen, M. B. The "low level" airman in retesting and basic training: A sociopsychological study. Journal of Social Psychology, 1961, 55, 177-190.

Jensen, M. B. Intelligence and learning in old age. Whipple, Arizona: Veterans Administration, 1966 (Mimeographed).

Jones, F. D. & Peters, H. N. An experimental evaluation of group psychotherapy. Journal of Abnormal & Social Psychology, 1952, 47, 345-353.

Jones, R. E. Personality changes in psychotics following prefrontal lobotomy. Journal of Abnormal and Social Psychology, 1949, 44, 315-328.

Joseph, A. & Murray, V. F. Chamorros and Carolinians of Saipan. Personality studies. Cambridge, Mass: Harvard University Press, 1951, 119-135, 329-336.

Judson, A. J. & MacCasland, B. W. The effects of chlorpromazine on psychological test scores. Journal of Consulting Psychology, 1960, 24, 192 (Abstract).

Kainer, R. K. The Porteus Maze test and the delay of gratification. (Doctoral dissertation, Columbia University) Ann Arbor, Mich.: University Microfilms, 1965. No. 65-13,960.

Karpeles, L. M. A further investigation of the Porteus Maze test as a discriminative measure in delinquency. Journal of Applied Psychology, 1932, 16, 427-437.

King, H. E. Intellectual Function; Chapter 14 in Mettler, F. A. (Ed.), Columbia-Greystone Associates - Selective partial ablation of the frontal cortex. New York: Paul B. Hoeber, Inc., 1949, 178-207.

Koskoff, Y. D., Dennis, W., Lazovik, D. and Wheeler, E. T. The psychological effects of frontal lobotomy performed for the alleviation of pain. Research Publications Association for Research in Nervous and Mental Disease, 1948, 27, 723-753.

Kunce, J., & Worley, B. Relationship of circular pencil mazes with cognitive and emotional distress. Journal of Clinical Psychology, 1966, 22 (3), 303-305.

La Barba, R. C. Relation of color responses on the Rorschach to qualtitative scores on the Porteus Maze test. Perceptual and Motor Skills, 1965, 21, 61-62.

Landis, C. Psychological observations on psychosurgery patients. Psychiatric Quarterly, 1951, 25, 409-417.

Landis, C. & Erlick, D. An analysis of the Porteus Maze test as affected by psychosurgery. American Journal of Psychology, 1950, 63, 557-566.

Landis, C., Garrison, M., Hamwi, V., Stauffer, A. K., King, H. E. and Young, K. M. Test results one year after operation. Chapter 23 in Mettler, F. A. (Ed.), Columbia-Greystone Associates-Selective partial ablation of the frontal cortex. New York: Paul B. Hoeber, Inc., 1949, 302-305.

Landis, C., Zubin, J. and Mettler, F. A. The functions of the human frontal lobe. Journal of Psychology, 1950, 30, 123-138.

Leibowitz, M. Effects of psychological support on test performance of four types of narcotic addicts. (Doctoral dissertation, New York University) Ann Arbor, Mich.: University Microfilms, 1966, No. 67-614.

Leiter, R. G. The Leiter standardized profile and a system of psychological nomenclature for mental illness. The Psychological Service Center Journal, 1959, 11, 79-87.

Loranger, A. W. & Misiak, H. Critical flicker frequency and some intellectual functions in old age. Journal of Gerontology, 1959, 14, 323-327.

Loranger, A. W. & Misiak, H. K. The performance of aged females on five non-language tests of intellectual functions. Journal of Clinical Psychology, 1960, 16, 189-191.

Louttit, C. M. Test performance of a selected group of part-Hawaiians. Journal of Applied Psychology, 1931, 15, 43-52.

Louttit, C . M. & Stackman, H. The relationship between Porteus Maze and Binet test performance. Journal of Educational Psychology, 1936, 27, 18-25.

Lubin, B., Wallis, R. & Paine, C. Patterns of psychological test usage in the United States: 1935-1969. Professional Psychology, 1971, 2, 70-74.

MacCrone, I. D. Preliminary results from the Porteus Maze tests applied to native school children. South African Journal of Science, 1928, 25, 481-484. (Psychological Abstracts, 1929, 3, 292, No. 2106)

Malmo, R. B. Psychological aspects of frontal gyrectomy and frontal lobotomy in mental patients. Research Publications Association for Research in Nervous and Mental Disease, 1948, 27, 537-564.

Matarazzo, J. D., Ulett, G. A. & Saslow, G. Human Maze performance as a function of increasing levels of anxiety. Journal of General Psychology, 1955, 53, 79-95.

Medina, R. F., Pearson, J. S. & Buchstein, H. F. The long term evaluation of prefrontal lobotomy in chronic psychotics. Journal of Nervous and Mental Diseases, 1954, 119, 23-30.

Meier, M. J. & French, L. A. Longitudinal assessment of intellectual functioning following unilateral temporal lobectomy. Journal of Clinical Psychology, 1966, 22, 22-27.

Meier, M. J. & Martin, W. E. Intellectual changes associated with Levodopa therapy. Journal of American Medical Association 1970, 213, 465-466.

Meier, M. J. & Okayama, M. Behavior assessment. Geriatrics, 1969, 24 (Nov) 95-110.

Meier, M. J. & Story, J. L. Selective impairment of Porteus Maze test performance after right sub-thalamotomy. Neuropsychologia, 1967, 5, 181-189.

Michaels, J. J. & Schilling, M. E. The correlations of the intelligence quotients of the Porteus Maze and Binet-Simon tests in two hundred neuropsychiatric patients. American Journal of Orthopsychiatry, 1936, 6, 71-74. (a)

Michaels, J. J. & Schilling, M. E. An attempt to determine the degree of antisocial behavior in psychopathic personalities and its correlation with the Porteus Maze and Binet-Simon tests. American Journal of Orthopsychiatry, 1936, 6, 397-405. (b)

Minski, L. & Desai, M. Aspects of personality in peptic ulcer patients. A comparison with hysterics. British Journal of Medical Psychology. 1955, 28, 113-134.

Morganthau, D. R. Some well-known mental tests evaluated and compared. Archives of Psychology. No. 52. New York: Archives of Psychology, Columbia University, 1922.

Myatt, M. F. A study of the relationship between motivation and test performance of patients in a rehabilitation ward. Doctoral dissertation, University of Minnesota, 1951.

O'Shea, H. E., Elsom, K. & Higbe, R. V. Studies of the B vitamins in the human subject. IV, Mental changes in experimental deficiency. American Journal of the Medical Sciences, 1942. 203, 388-397.

Perkins, N. L. Human reactions in a maze of fixed orientation. Comparative Psychology Monographs, 1927, 4, (21).

Peters, H. N. Circular pencil Maze performance in chronic schizophrenics. Journal of Clinical Psychology, 1956, 12, 170-173.

Peters, H. N., & Jones, F. D. Evaluation of group psychotherapy by means of performance tests. Journal of Consulting Psychology, 1951, 15, 363-367.

Peterson, J. & Telford, C. W. Results of group and of individual tests applied to the practically pure-blood Negro children on St. Helena Island. Journal of Comparative Psychology, 1930, 11, 115-144.

Petrie, A. Personality and the frontal lobes. London: Routledge & Kegan Paul, Ltd., 1952.

Petrie, A. & LeBeau, J. Psychologic changes in man after chlorpromazine and certain types of brain surgery. Journal of Clinical and Experimental Psychopathology, 1956, 17, 170-179.

Piddington, M. & Piddington, R. Report of field work in Northwestern Australia. Oceania, 1932, 2, 342-358.

Phillips, E. L., Berman, I. R. & Hanson, H. B. Intelligence and personality factors associated with poliomyelitis among school age children. Monographs of the Society for Research in Child Development, 1948, 12 (2).

Pomeroy, D. S. Ameliorative effects of "Counseling" upon maze performance following experimentally induced stress. American Psychologist, 1950, 5, 327. (Abstract).

Porteus, S. D. Mental tests for feeble-minded: A new series. Journal of Psycho-asthenics, 1915, 19, 200-213. (a)

Porteus, S. D. Motor-intellectual tests for mental defectives. Journal of Experimental Pedagogy, 1915, 3, 127-135. (b)

Porteus, S. D. Mental tests with delinquents and Australian aboriginal children. Psychological Review, 1917, 24, 32-42.

Porteus, S. D. The measurement of intelligence: Six hundred and fifty-three children examined by the Binet and Porteus tests. Journal of Educational Psychology, 1918, 9, 13-31.

Porteus, S. D. Porteus Tests - The Vineland Revision. Vineland, N. J.: Training School Research Publication, No. 16, 1919, 1-44.

Porteus, S. D. A study of the personality of defectives with a social ratings scale. Vineland, N. J.: Training School, Research Publication No. 23, 1920, 1-22.

Porteus, S. D. Studies in mental deviations. Vineland, N. J.: Smith Printing House, 1923.

Porteus, S. D. Guide to the Porteus Maze tests. Vineland, N. J.: Training School Publication, 1924.

Porteus, S. D. The psychology of a primitive people: A study of the Australian aborigine. London: Edward Arnold & Co., 1931.

Porteus, S. D. The Maze test and mental differences. Vineland, N. J.: Smith Printing and Publishing House, 1933.

Porteus, S. D. Primitive intelligence and environment. New York: MacMillan, 1937.

Porteus, S. D. Racial group differences in mentality. Tablae Biologicae, Den Haag, Holland, 1939, 18, 66-75. (Psychological Abstracts 1946, 20, 171.) (a)

Porteus, S. D. The validity of the Porteus Maze test. Journal of Educational Psychology, 1939, 30, 172-178. (b)

Porteus, S. D. Qualitative performance in the Maze test. Vineland, N. J.: Smith Printing House, 1942.

Porteus, S. D. Q-scores, temperament, and delinquency. Journal of Social Psychology, 1945, 21, 81-103.

Porteus, S. D. Porteus Maze tests. In Harriman, P. L. (Ed.). Encyclopedia of Psychology. New York: Philosophical Library, 1946, 537-544.

Porteus, S. D. The Porteus Maze test and intelligence. Palo Alto, California: Pacific Books, 1950.

Porteus, S. D. The Porteus Maze test manual. London: George G. Harrap & Co., 1952.

Porteus, S. D. Maze test qualitative aspects. British Journal of Medical Psychology, 1954, 27, 72-79.

Porteus, S. D. The Maze test: Recent advances. Palo Alto, California: Pacific Books, 1955.

Porteus, S. D. Maze test reactions after chlorpromazine. Journal of Consulting Psychology, 1957, 21, 15-21.

Porteus, S. D. What do the Maze tests measure? Australian Journal of Psychology, 1958, 10, 245-256.

Porteus, S. D. The Maze test and clinical psychology. Palo Alto, California: Pacific Books, 1959. (a)

Porteus, S. D. The Porteus Mazes. The Supplement Series of mazes. New York: The Psychological Corporation, 1959. (b)

Porteus, S. D. Ethnic group differences. Mankind Quarterly, 1961, 1, 187-200.

Porteus, S. D. Maze test reports. Perceptual and Motor Skills, 1962, 14, 58.

Porteus, S. D. Porteus Maze tests: Fifty years application. Palo Alto, California: Pacific Books, 1965.

Porteus, S. D. & Babcock, M. E. Temperament and race. Boston: Richard G. Badger, 1926.

Porteus, S. D., & Barclay, J. E. A further note on chlorpromazine: Maze reactions. Journal of Consulting Psychology, 1957, 21, 297-299.

Porteus, S. D. & Bassett, D. M. Sex differences in Porteus Maze test performance. Training School Bulletin, 1920, 22, 105-120.

Porteus, S. D., Bochner, S., Russell, J. & David, K. Age as a factor in Australid mentality. Perceptual and Motor Skills, 1967, 25, 3-16.

Porteus, S. D. & Ching, E. T. Research with feebleminded. Hawaii Medical Journal, 1959, 18, 491-493.

Porteus, S. D. & David, K. Australid mental development and geriatric decline. Perceptual and Motor Skills, 1966, 23, 75-87.

Porteus, S. D., Dewey, D. M. & Bernreuter, R. G. Race and social differences in performance tests. Genetic Psychology Monographs, 1930, 8, 93-208.

Porteus, S. D., & Gregor, A. J. Studies in intercultural testing. Perceptual and Motor Skills, 1963, 16, 705-724.

Porteus, S. D. & Kepner, R. D. Mental changes after bilateral prefrontal lobotomy. Genetic Psychology Monographs, 1944, 29, 3-115.

Porteus, S. D. & Peters, H. N. Maze test validation and psychosurgery. Genetic Psychology Monographs, 1947, 36, 3-86.

Poull, L. E. & Montgomery, R. P. The Porteus Maze test as a discriminative measure in delinquency. Journal of Applied Psychology, 1929, 13, 145-151.

Purcell, K. A note on Porteus Maze and Wechsler-Bellevue scores as related to antisocial behavior. Journal of Consulting Psychology, 1956, 20, 361-364.

Purcell, K., Turnbull, J. W. & Bernstein, L. Distinctions between subgroups of asthmatic children: Psychological test and behavior rating comparisons. Journal of Psychosomatic Research, 1962, 6, 283-291.

Rafi, A. A. Motor performance of certain categories of mental patients. Perceptual and Motor Skills, 1960, 10, 39-42.

Ray, P. C. Maze test performance of the Bhil of central India. Bulletin of the Department of Anthropology, 1953, 2, 83-90. Cited by Porteus, S. D. and Gregor, A. J. Studies in intercultural testing. Perceptual and Motor Skills, 1963, 16, 705-724.

A report of intelligence testing by the Porteus Maze test. Education of Greater Shanghai, 1935, 2, No. 8 (Special Number). (Psychological Abstracts, 1936, 10, 581, No. 517)

Roberts, A. H. & Erikson, R. V. Delay of gratification, Porteus Maze test performance, and behavioral adjustment in a delinquent group. Journal of Abnormal Psychology, 1968, 73, 449-553.

Robinson, M. F. What price lobotomy? Journal of Abnormal and Social Psychology, 1946, 41, 421-436.

Robinson, M. F. Personality change after psychosurgery. Chapter 17 in Freeman, W. & Watts, J. W., Psychosurgery (2nd ed.). Springfield, Ill.: Charles C. Thomas, 1950, 314-328.

Ross, E. L. S. Vocational tests for mental defectives. Studies in Mental inefficiency, 1921, 2, 1-6.

Ruebush, B. K., Byrum, M. & Farnham, L. J. Problem solving as a function of children's defensiveness and parental behavior. Journal of Abnormal and Social Psychology, 1963, 67, 355-362.

Sanderson, M. H. Performance of fifth, eighth, and eleventh grade children in the Porteus Qualitative Maze tests. Journal of Genetic Psychology, 1945, 67, 57-65.

Sarason, S. B., Davidson, K. S., Lighthall, F. F., Waite, R. R. & Ruebush, B. K. Anxiety in elementary school children; a report of research. New York: John Wiley & Sons, Inc., 1960, 171-188.

Satter, G. Retarded adults who have developed beyond expectation: Part III: Further analysis and summary. Training School Bulletin, 1955, 51, 237-243.

Satter, G. & McGee, E. Retarded adults who have developed beyond expectation: Part I: Intellectual functions. Training School Bulletin, 1954, 51, 43-55.

Schalling, D. & Rosén, A. Porteus Maze differences between psychopathic and non-psychopathic criminals. British Journal of Social and Clinical Psychology, 1968, 7, 224-228.

Scott, T. The one-error per trial method of maze learning. Psychological Bulletin, 1936, 33, 768-769. (Abstract)

Shakow, D. & Millard, M. S. A psychometric study of 150 adult delinquents. Journal of Social Psychology, 1935, 6, 437-457.

Shapiro, M. B., Campbell, D., Harris, A. & Dewsbery, J. P. Effects of E.C.T. upon psychomotor speed and the "distraction effect" in depressed psychiatric patients. Journal of Mental Science, 1958, 104, 681-695.

Shapiro, M. B., Kessell, R. & Maxwell, A. E. Speed and quality of psychomotor performance in psychiatric patients. Journal of Clinical Psychology, 1960, 16, 266-271.

Shapiro, M. B., Slater, P. & Campbell, D. The effects of distraction on psychomotor slowness in cooperative depressed, and schizophrenic subjects. British Journal of Social and Clinical Psychology, 1962, 1, 121-126.

Sheer, D. E. Psychometric studies. In Lewis, N. D. C., Landis, C. & King, H. E. (Eds.), Studies in topectomy. New York: Grune & Stratton, 1956, 56-74.

Sheer, D. E., Shuttleworth, M., Stopol, H. B., Turney, J. F. & Hamwi, V. Psychometric studies. In Mettler, F. A. (Ed.), Columbia-Greystone Associates — Psychosurgical Problems. Philadelphia: The Blakiston Company, 1952.

Small, K. Planning as a non-intellective component of intelligent behavior. (Doctoral dissertation, Columbia University) Ann Arbor, Michigan: University Microfilms, 1954, No. 8833.

Smith, A. Changes in Porteus Maze scores of brain-operated schizophrenics after an eight-year interval. Journal of Mental Science, 1960, 106, 967-978.

Smith, A. & Kinder, E. F. Changes in psychological test performances of brain-operated schizophrenics after eight years. Science, 1959, 129, 149-150.

Smith, F. O. Differential reactions of human beings in the maze. Pedagogical Seminary. (Journal of Genetic Psychology) 1927, 34, 394-405.

Sparling, M. E. Intelligence of children - the relationship between Binet and Porteus scores. American Journal of Mental Deficiency, 1941, 46, 60-62.

Sternberg, A. L. The relationship between manifest anxiety and stress, and qualitative performance on the Porteus Maze test. Unpublished master's thesis, New Mexico Highlands University, 1964.

Sterne, D. M. The Benton, Porteus and WAIS Digit Span tests with normal and brain-injured subjects. Journal of Clinical Psychology, 1969, 25, 173-175.

Stotsky, B. A., Sacks, J. M. & Daston, P. Predicting the work performance of psychiatric aides by psychological tests. Journal of Counseling Psychology, 1956, 3, 193-199.

Sundberg, N. D. The practice of psychological testing in clinical services in the United States. American Psychologist, 1961, 16, 79-83.

Sydow, D. W. A psychometric differentiation between functional psychotics and non-psychotics with organic brain damage. Doctoral dissertation, University of Minnesota, 1953.

Tachibana, K. Experimental study of Maze-learning of the aged. Japanese Journal of Psychology, 1931, 6, 207-222.

Telford, C . W. Test performance of full and mixed-blood North Dakota Indians. Journal of Comparative Psychology, 1932, 14, 123-145.

Tizard, J. The Porteus Maze test and intelligence: A critical survey. British Journal of Educational Psychology, 1951, 21, 172-185.

Tizard, J., O'Connor, N. & Crawford, J. M. The abilities of adolescent and adult high-grade male defectives. Journal of Mental Science, 1950, 96, 889-907.

Tobias, J. & Gorelick, J. The Porteus Maze test and the appraisal of retarded adults. American Journal of Mental Deficiency. 1962, 66, 600-606.

Tow, P. M. Personality changes following frontal leucotomy. London: Oxford University Press, 1955, 66-68, 171-181, 222, 241, 242.

Tsao, J. C. A note on the Porteus Maze test. Chinese Journal of Psychology, 1937, 1, 252-264. (Psychological Abstracts, 1937, 11, 677-678, No. 5994).

Vernon, P. E. A study of the norms and the validity of certain mental tests at a child guidance clinic, Part II. British Journal of Educational Psychology, 1937, 7, 115-137.

Verrill, B. V. An investigation of the concept of impulsivity. (Doctoral dissertation, University of Houston) Ann Arbor, Michigan: University Microfilms, 1958, No. 58-2009.

Vicary, T. C. An investigation into the use of non-verbal tests of intelligence in India. Unpublished master's thesis, University of London, 1938. Cited by Porteus, S. D. and Gregor, A. J. Studies in intercultural testing. Perceptual and Motor Skills, 1963, 16, 705-724.

Waite, J. H. & Neilson, I. L. A study of effects of hookworm infection upon the mental development of North Queensland school children. Medical Journal of Australia, 1919, 1, 1-8.

Weintraub, S. A. Cognitive and behavioral impulsivity in internalizing, externalizing and normal children. Doctoral dissertation, University of Minnesota, 1968.

Weisenburg, T. & McBride, K. Aphasia. A clinical and psychological study. New York: Commonwealth Fund, 1935.

Weisenburg, T., Roe, A. & McBride, K. E. Adult intelligence, a psychological study of test performances. New York: Commonwealth Fund, 1936.

Wittenborn, J. R. & Mettler, F. A. Some psychological changes following psychosurgery. Journal of Abnormal and Social Psychology, 1951, 46, 548-556.

Worthington, M. R. A study of some commonly used performance tests. Journal of Applied Psychology, 1926, 10, 216-227.

Wright, C. The qualitative performance of delinquent boys on the Porteus Maze test. Journal of Consulting Psychology, 1944, 8, 24-26.

Zeckel, A. & Van Der Kolk, J. J. A comparative intelligence test of groups of children born deaf and of good hearing, by means of the Porteus test. American Annals of the Deaf, 1939, 84, 114-123.

Zubin, J., Sutton, S., Salzinger, K., Salzinger, S., Burdock, E. I. & Peretz, D. A biometric approach to prognosis in schizophrenia. In Hoch, P. H. & Zubin, J. (Eds.), Comparative epidemiology of the mental disorders. New York: Grune & Stratton, 1961. 143-203.

Figure captions

Figure 1 Lateral view of the left hemisphere of the brain showing the site of anterior and posterior operations with respect to Brodmann's areas.

Figure 2 Mazes from the first published series, published in the Journal of Psychoasthenics (Porteus, 1915 a).

Figure 3 Mazes from the first published series, published in the Journal of Psychoasthenics (Porteus, 1915 a).

Figure 4 Mazes published in Porteus tests--The Vineland Revision (Porteus, 1919).

Figure 5 Maze published in Porteus tests--The Vineland Revision (Porteus, 1919).

Figure 6 Maze published in Porteus tests--The Vineland Revision (Porteus, 1919).

Figure 7 Maze published in Porteus tests--The Vineland Revision (Porteus, 1919).

Figure 8 Maze published in Guide to the Porteus Maze test (Porteus, 1924).

Figure 9 Maze published in Guide to the Porteus Maze test (Porteus, 1924).

Figure 10 Mazes published in The Maze test and mental differences (Porteus, 1933).

Figure 11 Maze published in The Maze test and mental differences (Porteus, 1933).

Figure 12 Mazes published in The Maze test and mental differences (Porteus, 1933).

Figure 13 Maze published in The Maze test and mental differences (Porteus, 1933).

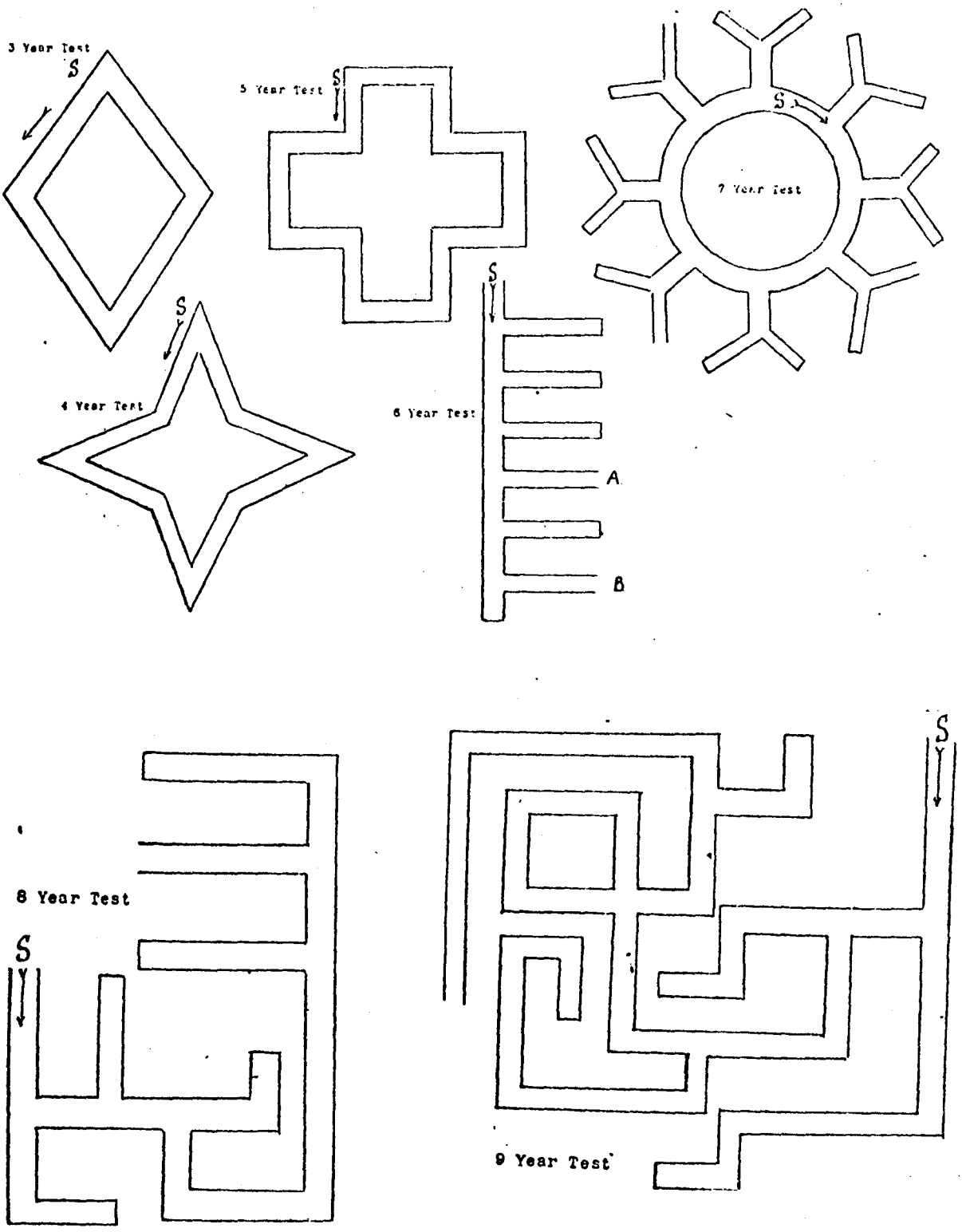


Figure 2

Mazes from the first published series, published in the Journal of Psychoasthenics (Porteus, 1915a).

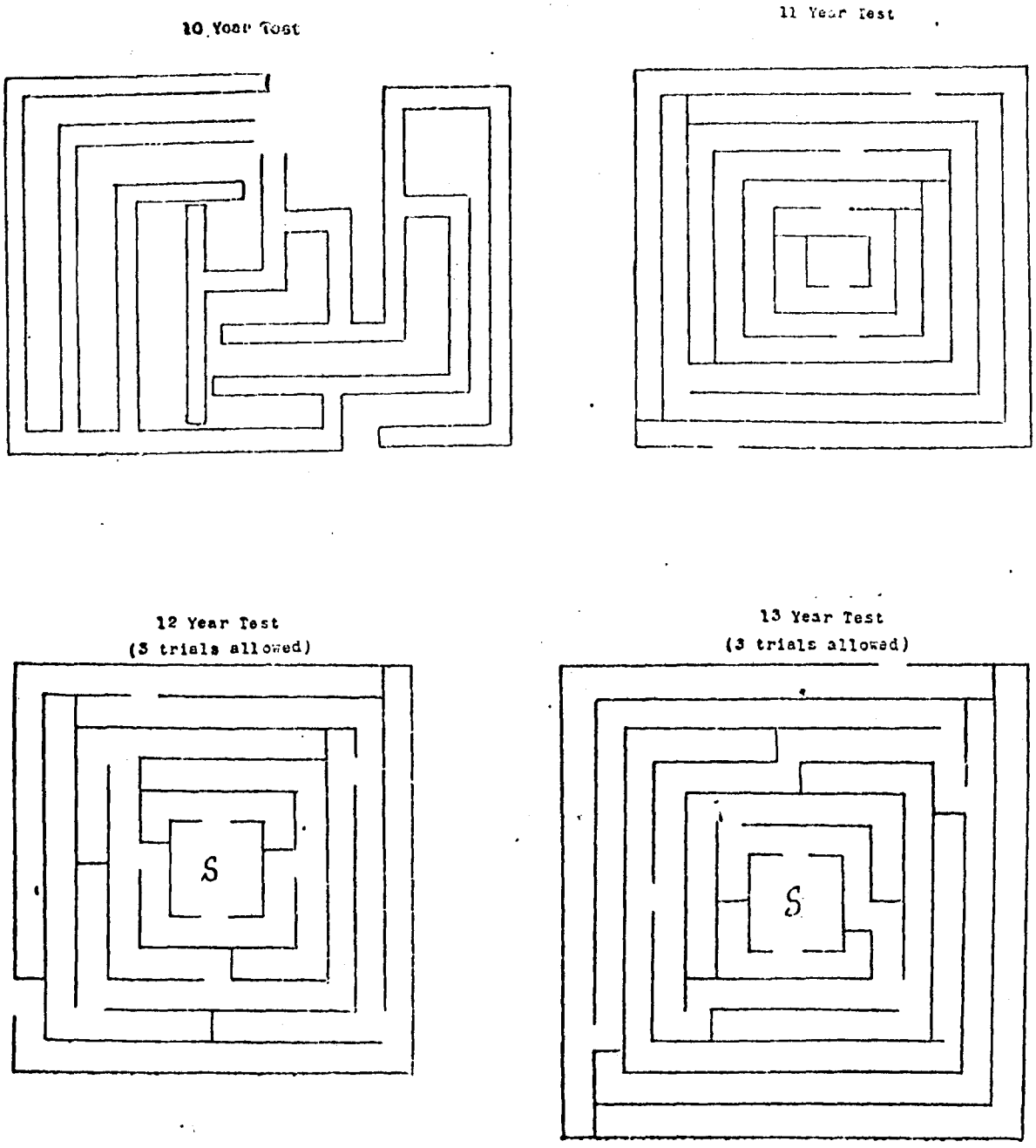
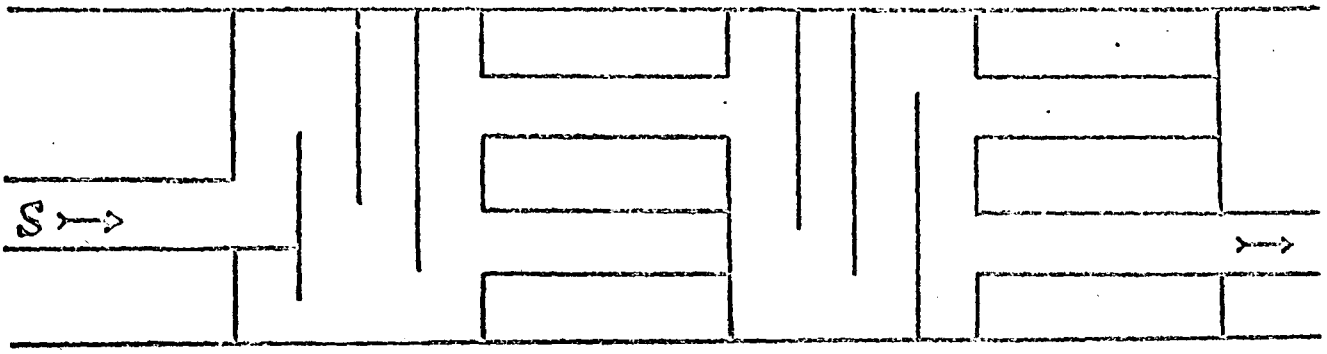
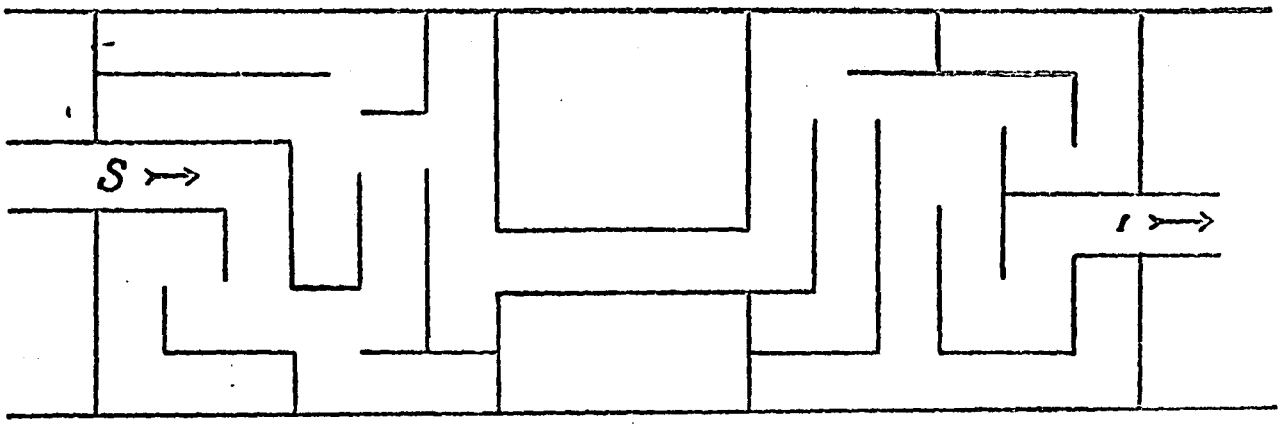


Figure 3

Mazes from the first published series, published in the Journal of Psychoasthenics (Porteus, 1915a).



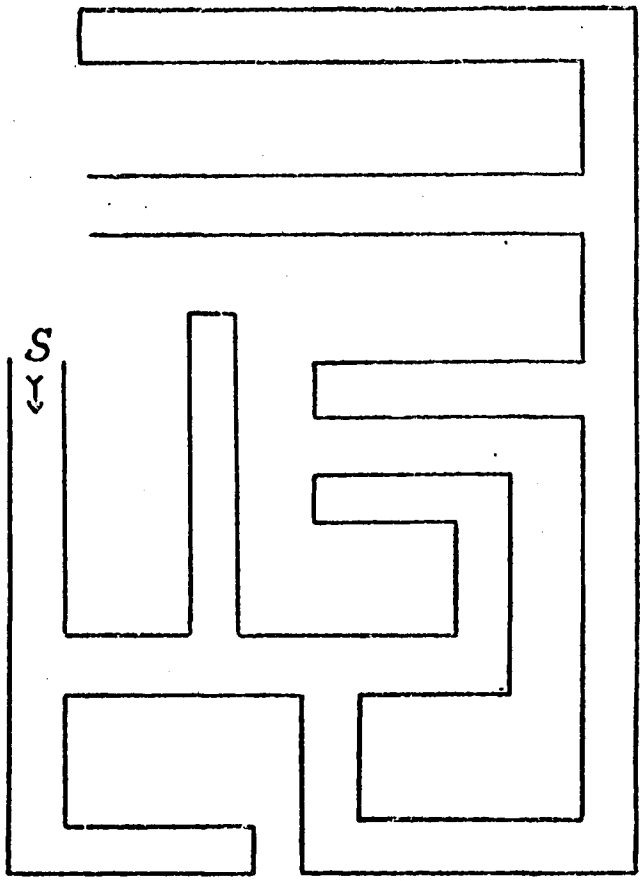
Porteus Test—Vineland Revision Year VI



Porteus Test—Vineland Revision Year VII

Figure 4

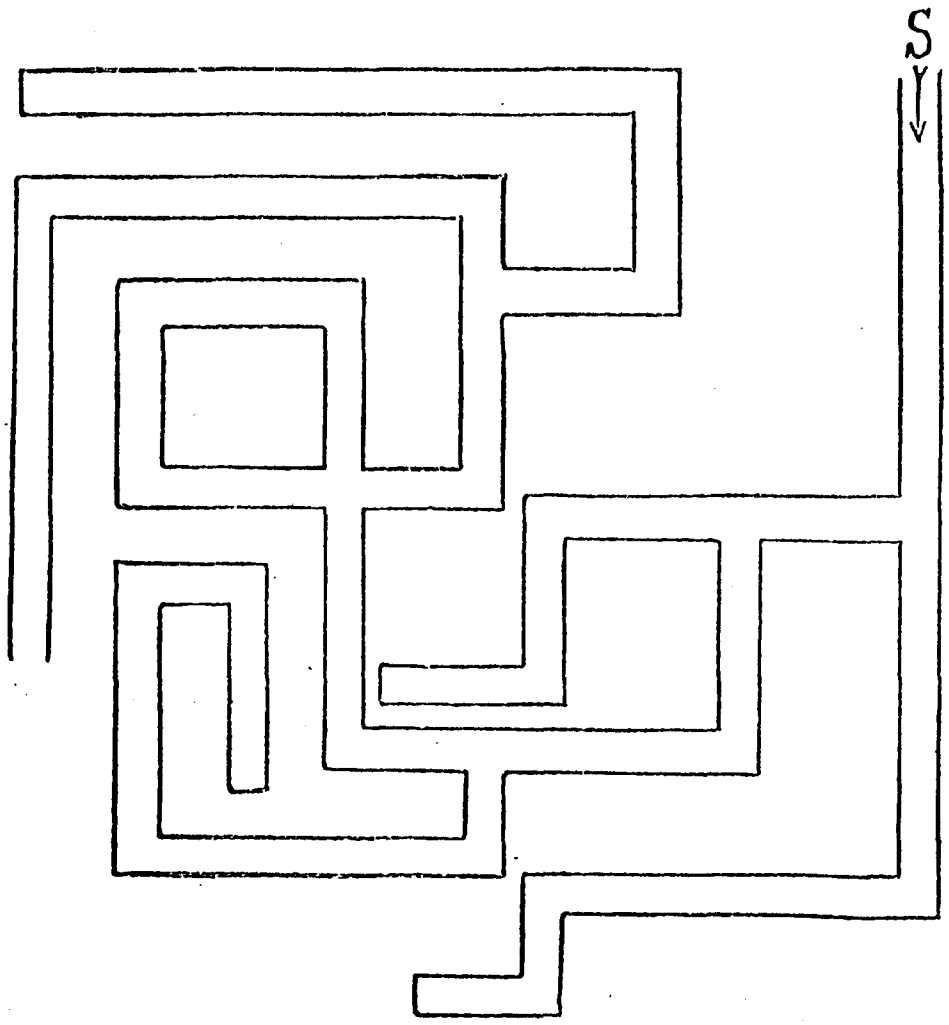
Mazes published in Porteus tests—The Vineland Revision (Porteus, 1919).



Porteus Test—Vineland Revision Year VIII

Figure 5

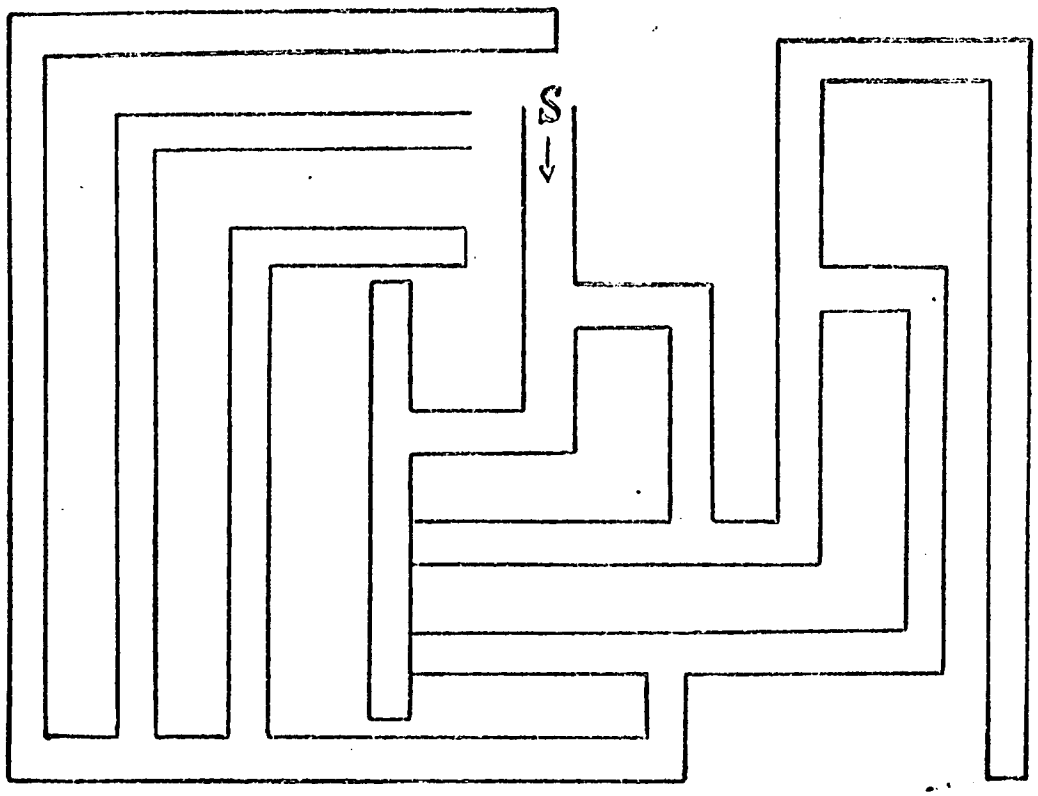
Maze published in Porteus tests—The Vineland Revision (Porteus, 1919).



Porteus Test—Vineland Revision Year IX

Figure 6

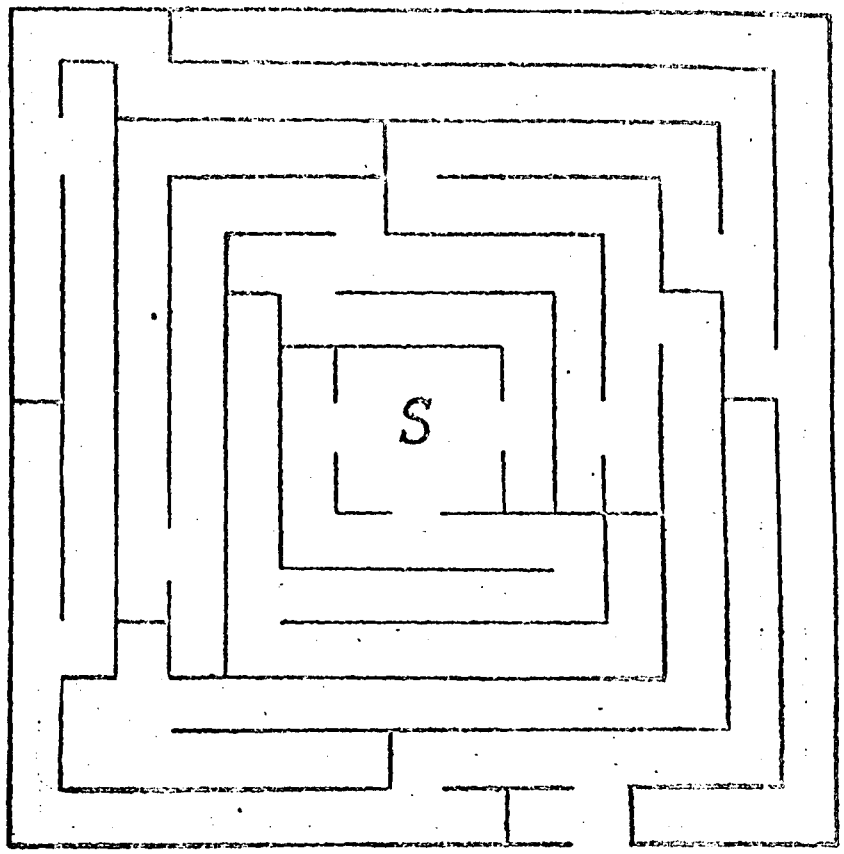
Maze published in Porteus tests—The Vineland Revision (Porteus, 1919).



Porteus Test—Vineland Revision Year X

Figure 7

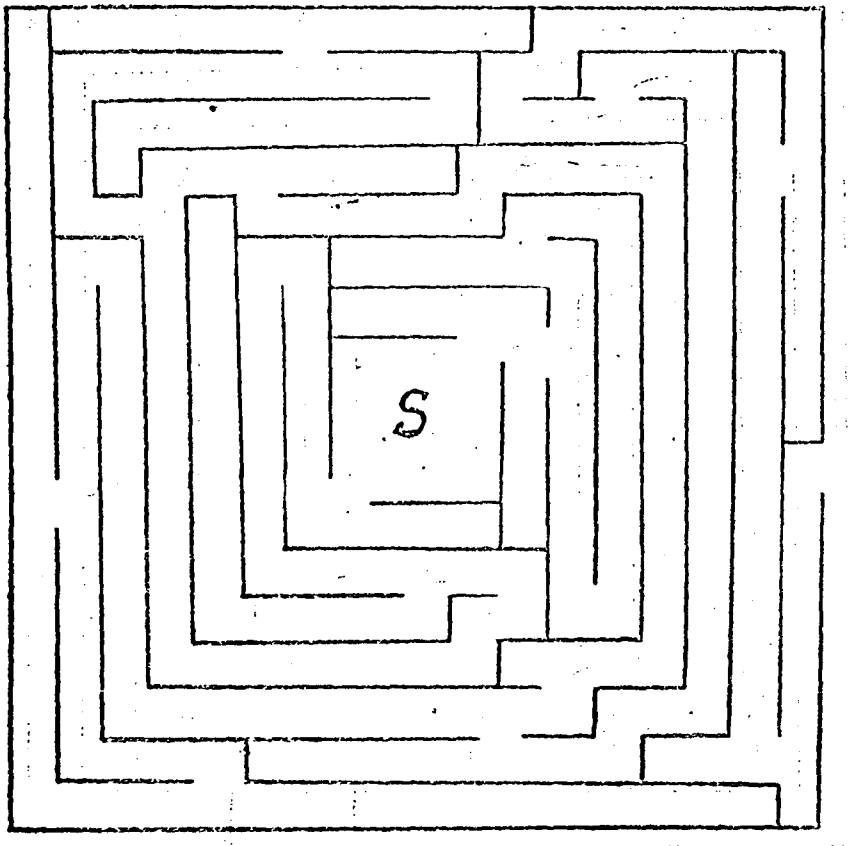
Maze published in Porteus tests—The Vineland Revision (Porteus, 1919).



Porteus Test—Vineland Revision Adult I

Figure 8

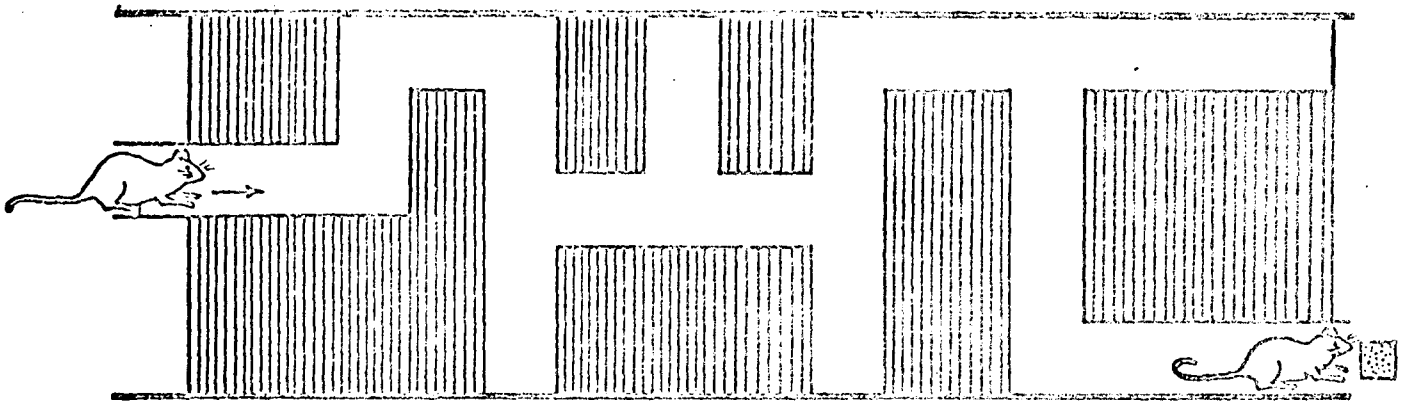
Maze published in Guide to the Porteus Maze test (Porteus, 1924).



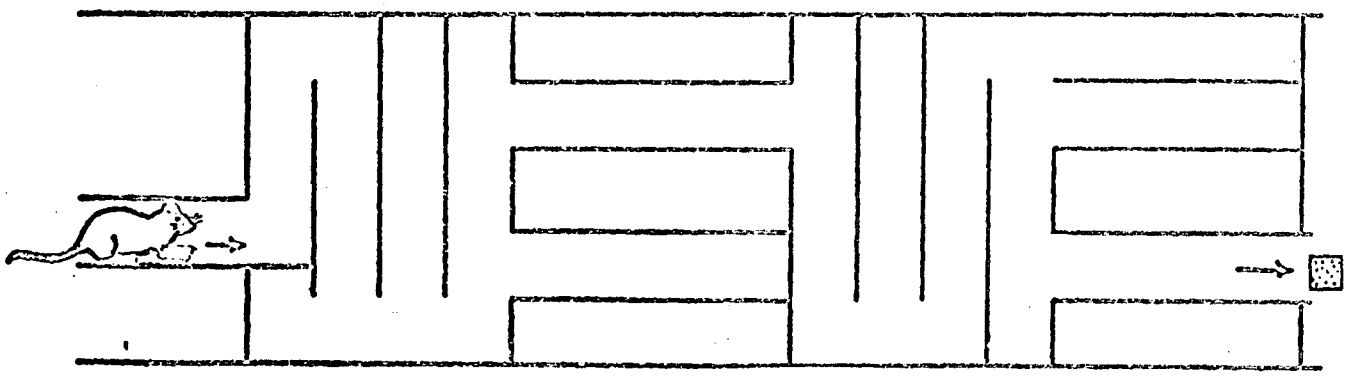
Porteus Test—Vineland Revision Adult II

Figure 9

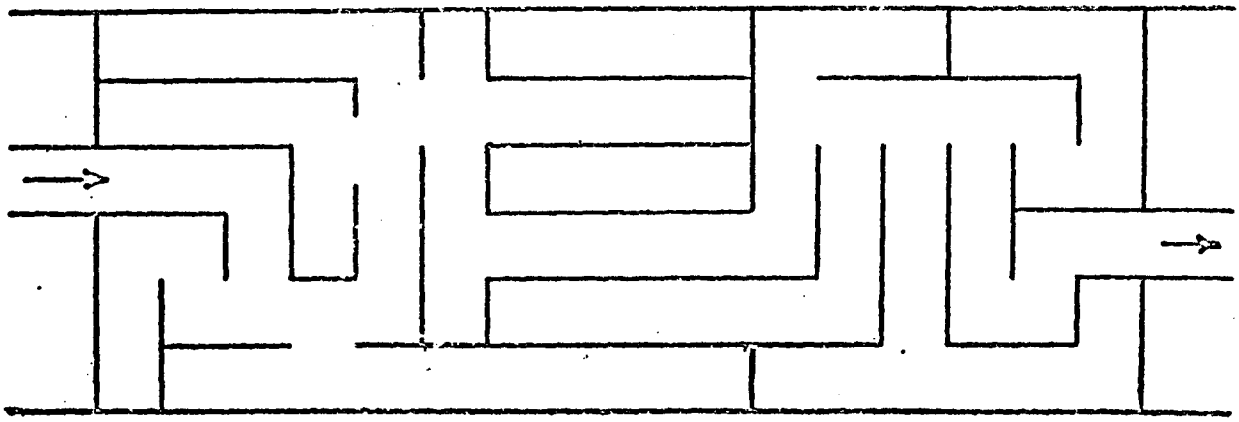
Maze published in Guide to the Porteus Maze test (Porteus, 1924).



Porteus Tests—Vineland Revision Year V

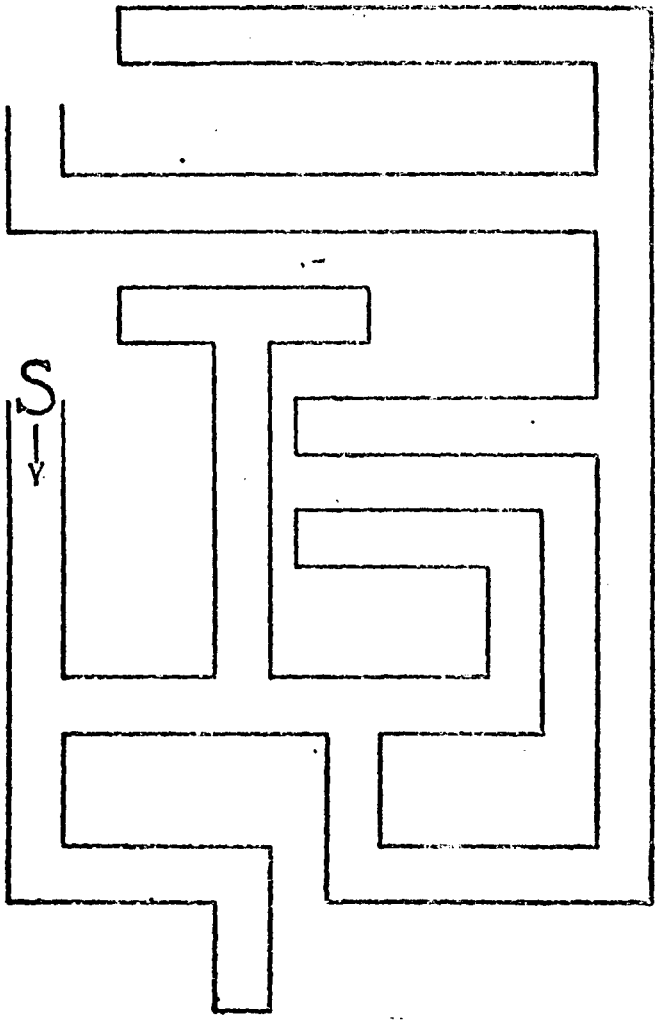


Porteus Tests—Vineland Revision Year VI



Porteus Tests—Vineland Revision Year VII

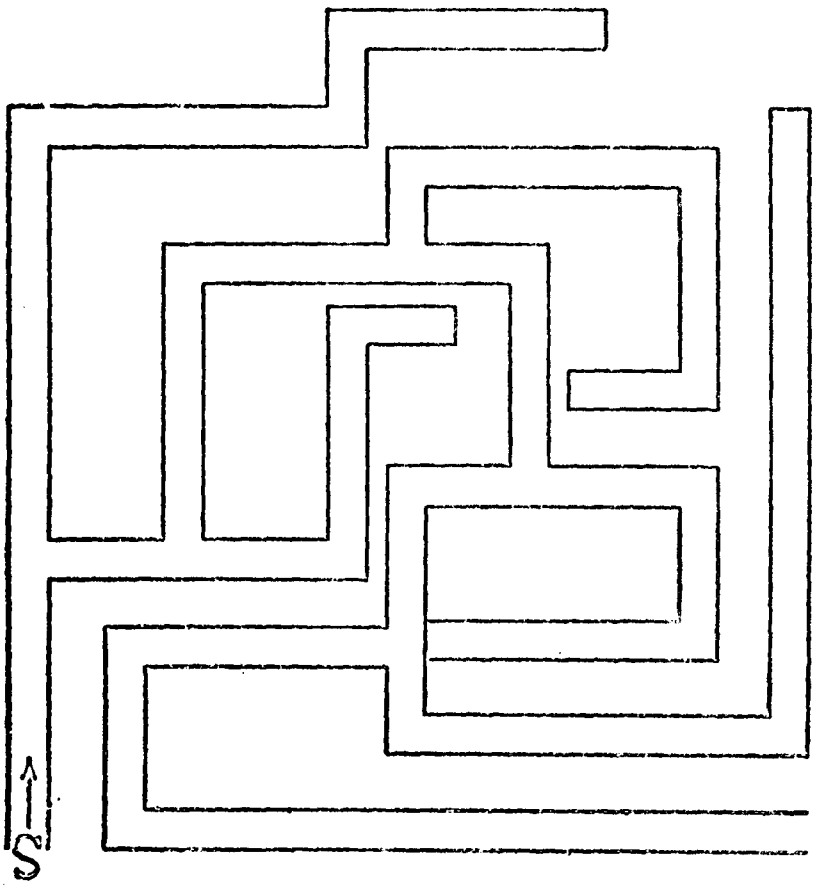
Figure 10
Mazes published in The Maze test and mental differences (Porteus, 1933).



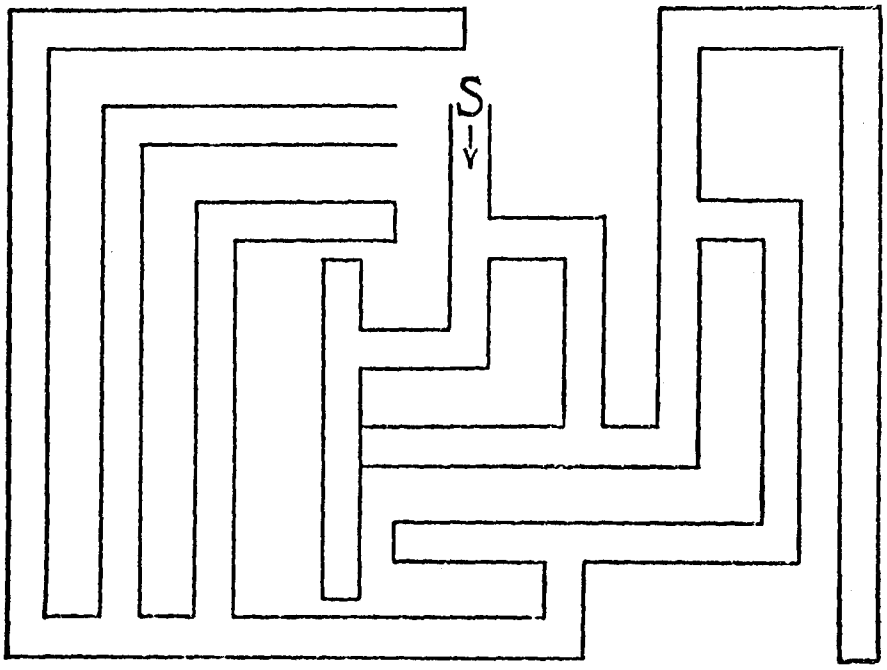
Porteus Tests--Vineland Revision Year VIII

Figure 11

Maze published in The Maze test and mental differences (Porteus, 1933).

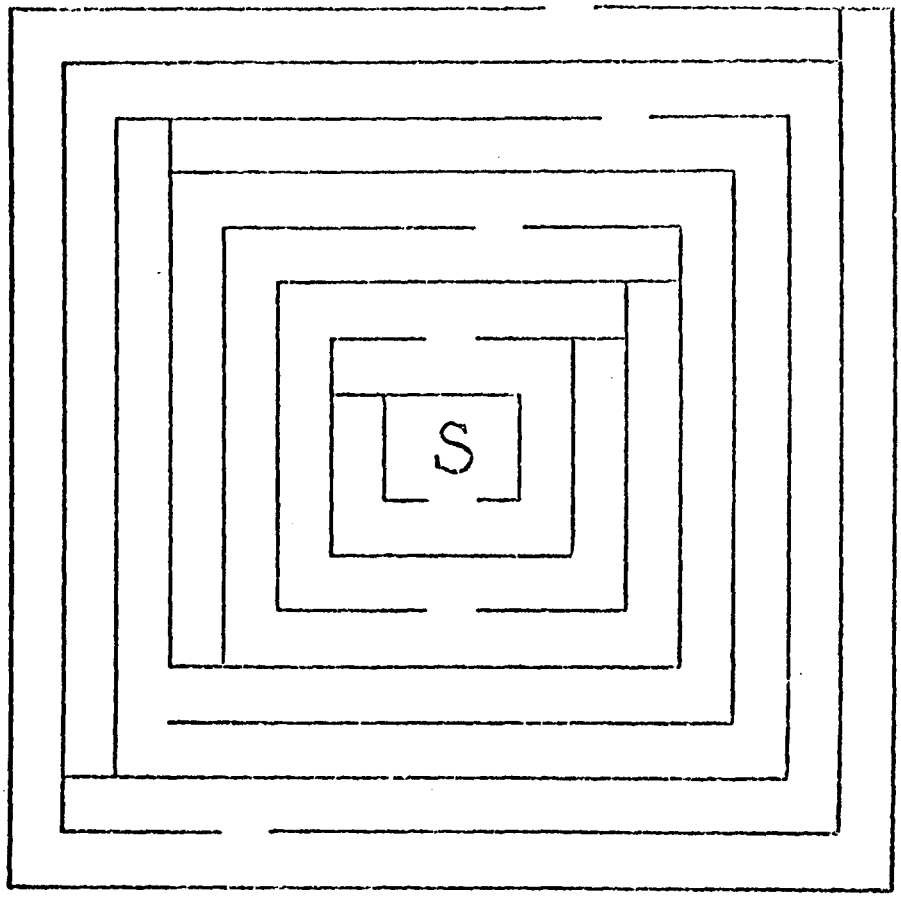


Porteus Tests—Vineland Revision Year IX



Porteus Tests—Vineland Revision Year X

Figure 12
Mazes published in The Maze test and mental differences (Porteus, 1933).



Porteus Tests—Vineland Revision Year XI

Figure 13

Maze published in The Maze test and mental differences (Porteus, 1933).