THE
THIRTY-NINTH YEARBOOK

OF THE

NATIONAL SOCIETY FOR THE STUDY OF EDUCATION

INTELLIGENCE: ITS NATURE AND NURTURE
PART I
COMPARATIVE AND CRITICAL EXPOSITION

Prepared by the Society's Committee

Leonard Carmichael, Frank N. Freeman, Florence L. Goodenough, Leta S. Hollingworth, Harold E. Jones, George D. Stoddard (Chairman), Lewis M. Terman, and Beth L. Wellman

Assisted by Members of the Society and Others

Edited by

GUY MONTROSE WHIPPLE

This Yearbook Will Be Discussed at the St. Louis Meeting of the Society, Saturday, February 24, 1940, 8:00 p.m., and Monday, February 26, 1940, 2:15 p.m.

PUBLIC SCHOOL PUBLISHING COMPANY
BLOOMINGTON, ILLINOIS
1940

Agents
PUBLIC SCHOOL PUBLISHING COMPANY
Bloomington, Illinois
PUBLISHERS OF ALL THE YEARBOOKS OF THE SOCIETY

The responsibilities of the Board of Directors of the National Society for the Study of Education in the case of yearbooks prepared by the Society's committees are (1) to select the subjects to be investigated, (2) to appoint committees calculated in their personnel to ensure consideration of all significant points of view, (3) to provide appropriate subsidies for necessary expenses, (4) to publish and distribute the committees' reports, and (5) to arrange for their discussion at the annual meetings.

The responsibility of the Yearbook Editor is to prepare the submitted manuscripts for publication in accordance with the principles and regulations approved by the Board of Directors in the "Guide for Contributors."

Neither the Board of Directors, nor the Yearbook Editor, nor the Society is responsible for the conclusions reached or the opinions expressed by the Society's yearbook committees.

Published February, 1940
First Printing, 3,000 copies

Printed by
The Plimpton Press
Norwood, Massachusetts

OFFICERS OF THE SOCIETY

FOR 1939–1940

Board of Directors

(Term of office expires March 1 of the year indicated)

WILLIAM C. BAGLEY (1942)
Columbia University, New York, New York

LEO J. BRUECKNER (1941)
University of Minnesota, Minneapolis, Minnesota

FRANK N. FREEMAN (1943) *
University of California, Berkeley, California

BESS GOODYKOONTZ, *Chairman* (1941)
Office of Education, Washington, D. C.

ERNEST HORN (1940)
State University of Iowa, Iowa City, Iowa

GRAYSON N. KEFAUVER (1942)
Stanford University, Palo Alto, California

RALPH W. TYLER (1943) †
University of Chicago, Chicago, Illinois

GUY MONTROSE WHIPPLE (*Ex-officio*)
Clifton, Massachusetts

Secretary-Treasurer

GUY MONTROSE WHIPPLE (1941)
Clifton, Massachusetts

* Elected for three years, beginning March 1, 1940.
† Reëlected for three years, beginning March 1, 1940.

THE SOCIETY'S COMMITTEE ON INTELLIGENCE: ITS NATURE AND NURTURE

LEONARD CARMICHAEL, Director of the Research Laboratory of Sensory Psychology and Physiology, and President, Tufts College, Medford, Massachusetts.

FRANK N. FREEMAN, Dean of Education, University of California, Berkeley, California.

FLORENCE L. GOODENOUGH, Research Professor, Institute of Child Welfare, University of Minnesota, Minneapolis, Minnesota.

* LETA S. HOLLINGWORTH, Professor of Education, Teachers College, Columbia University, New York, New York.

HAROLD E. JONES, Director of the Institute of Child Welfare, and Professor of Psychology, University of California, Berkeley, California.

GEORGE D. STODDARD (*Chairman*), Director of the Iowa Child Welfare Research Station, and Dean of the Graduate College, State University of Iowa, Iowa City, Iowa.

LEWIS M. TERMAN, Professor of Psychology, Stanford University, Palo Alto, California.

BETH L. WELLMAN, Professor of Child Psychology, Iowa Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

---

* Deceased November 27, 1939.

ASSOCIATED CONTRIBUTORS *

JOHN E. ANDERSON, Director, Institute of Child Welfare, University of Minnesota, Minneapolis, Minnesota.

L. DEWEY ANDERSON, Director of Psychological Research, Developmental Health Inquiry, Western Reserve University, Cleveland, Ohio.

FRANCES P. BARLOW, Utah State Agricultural College, Logan, Utah.

NANCY BAYLEY, Research Associate, Institute of Child Welfare, University of California, Berkeley, California.

GRACE E. BIRD, Professor of Educational Psychology, Rhode Island College of Education, Providence, Rhode Island.

BARBARA S. BURKS, Research Associate, Carnegie Institution of Washington, Cold Spring Harbor, Long Island.

HAROLD D. CARTER, Research Associate, Institute of Child Welfare, University of California, Berkeley, California.

R. B. CATTELL, G. Stanley Hall Professor of Genetic Psychology, Clark University, Worcester, Massachusetts.

HERBERT S. CONRAD, Research Associate, Institute of Child Welfare, University of California, Berkeley, California.

EDGAR A. DOLL, Director of Research, The Training School, Vineland, New Jersey.

M. R. HARROWER ERICKSON, Montreal Neurological Institute, McGill University, Montreal, Canada.

CECILE WHITE FLEMMING, Horace Mann School, New York, New York.

ARDEN FRANDSEN, Professor of Psychology, Utah State Agricultural College, Logan, Utah.

ARNOLD GESELL, Director of the Clinic of Child Development, Yale University, New Haven, Connecticut.

GERTRUDE HILDRETH, Psychologist, Lincoln School, Columbia University, New York, New York.

RALPH T. HINTON, JR., Mooseheart Laboratory for Child Research, Mooseheart, Illinois.

MARJORIE PYLES HONZIK, Research Associate, Institute of Child Welfare, University of California, Berkeley, California.

KATHRYN M. HORAN, Public School 500, Manhattan, Speyer School, New York, New York.

BYRON O. HUGHES, Research Associate in Child Development, University of Michigan, Ann Arbor, Michigan.

ADA P. JORGENSEN, Institute of Child Welfare, University of California, Berkeley, California.

* Includes contributors both to Part I and to Part II.

NEWELL C. KEPHART, Mental Hygienist, Wayne County Training School, Northville, Michigan.

JOHN R. KNOTT, Research Associate in Psychology, State University of Iowa, Iowa City, Iowa.

GEORGE KREEZER, Assistant Professor of Psychology, Cornell University, Ithaca, New York.

G. M. KUZNETS, formerly Instructor in Psychology, Stanford University, Palo Alto, California.

EDNA E. LAMSON, State Teachers College, Jersey City, New Jersey.

KURT LEWIN, Professor of Child Psychology, Iowa Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

DONALD B. LINDSLEY, Director, Psychological Clinic and Neurophysiological Laboratory, Bradley Home, East Providence, Rhode Island.

JANE LOEVINGER, Teaching Assistant, Department of Psychology, University of California, Berkeley, California.

IRVING LORGE, Research Associate, Institute of Educational Research, Columbia University, New York, New York.

FREDERICK H. LUND, Professor of Psychology, Temple University, Philadelphia, Pennsylvania.

OLGA MCNEMAR, Stanford University, Palo Alto, California.

KATHERINE MAURER, Psychometrist, Institute of Child Welfare, University of Minnesota, Minneapolis, Minnesota.

MELITA ODEN, Stanford University, Palo Alto, California.

WILLARD C. OLSON, Director of Research in Child Development, University of Michigan, Ann Arbor, Michigan.

MIRIAM C. PRITCHARD, Public School 500, Manhattan, Speyer School, New York, New York.

MARTIN L. REYMERT, Director, Mooseheart Laboratory for Child Research, Mooseheart, Illinois.

KATHERINE E. ROBERTS, Department of Psychology, Merrill-Palmer School, Detroit, Michigan.

ROBERT R. RUSK, Director, Scottish Council for Research in Education, Edinburgh, Scotland.

NATHAN W. SHOCK, Assistant Professor of Physiology, Medical School, University of California, Berkeley, California.

FRANK K. SHUTTLEWORTH, College of the City of New York, New York, New York.

HAROLD M. SKEELS, Assistant Professor of Psychology, Iowa Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

George S. Speer, Psychologist, Children's Service League, Springfield, Illinois.

S. D. Shirley Spragg, Instructor in Psychology, Barnard College, Columbia University, New York, New York.

Margaret Stanger, Ethical Culture School, New York, New York.

Elizabeth K. Starkweather, Merrill-Palmer School, Detroit, Michigan.

Mary Elizabeth Stippich, Assistant, Institute of Child Welfare, University of Minnesota, Minneapolis, Minnesota.

Robert L. Thorndike, Assistant Professor of Educational Psychology, Teachers College, Columbia University, New York, New York.

Robert C. Tryon, Associate Professor of Psychology, University of California, Berkeley, California.

William H. Voas, Psychologist, Winnetka Public Schools, Winnetka, Illinois.

Harold M. Williams, Supervisor for the Mentally Handicapped, State Department of Public Instruction, Madison, Wisconsin.

Paul A. Witty, Professor of Education, Northwestern University, Evanston, Illinois.

TABLE OF CONTENTS

FOR

PART I. COMPARATIVE AND CRITICAL EXPOSITION

SECTION III.— THE PHYSIOLOGY OF INTELLIGENCE

EDITOR'S PREFACE

In 1922 the Society published the *Twenty-First Yearbook,* entitled " Intelligence Tests and Their Use," and in 1928, the *Twenty-Seventh Yearbook,* entitled " Nature and Nurture "; Part I, " Their Influence upon Intelligence," and Part II, " Their Influence upon Achievement." In other yearbooks, as, for example, those dealing with " The Measurement of Educational Products," " The Education of Gifted Children," " Adapting the Schools to Individual Differences," " Educational Diagnosis," " The Grouping of Pupils," and " Child Development and the Curriculum," more or less extensive discussion is to be found concerning the nature and use of tests of intelligence and concerning the relative contributions of heredity and of environment to the making of adult mentality. On these accounts, accordingly, the present yearbook is not a first excursion for this Society into a *terra incognita,* however obscure the terrain and its boundaries may appear to remain after the current volume has been exhaustively inspected for guidance.

More particularly this *Thirty-Ninth Yearbook* is to be regarded as sequential to the *Twenty-Seventh Yearbook.* Early in the discussion of it by the Board of Directors it was declared that this " Stoddard Yearbook " was intended to be " a more positive study " than the " Terman Yearbook," and " not so much a debate upon controversial issues as an exploration of possibilities." But yearbooks have a way of developing their own issues and of thrusting forward divergencies of procedure and interpretation that no keenly intelligent committee can evade without stultifying their professional consciences. It is the editor's impression at least that the discussion of nature and nurture in 1940 is just as controversial as it was in 1928 — more so, indeed. So much the more is credit due the Chairman and the members of the Committee for their very frank and straightforward criticism of one another's methods, results, and interpretations set forth without personalities or rancor. They have felt that the problems they were treating were of the utmost import not only for psychological theory but also for social and educational practice.

That leads me to say that when the Directors expressed the expectation that this Yearbook should take the form of " a more positive study " and an " exploration of possibilities," it was urged particularly that there should be "a clear statement of the social implications of our present knowledge of nature and nurture," and that some space might well be devoted "to the implications for school organization." The Yearbook Committee will not need to invent reasons to explain why, in the text that follows, relatively little has been said, or at least said with unanimity, on these matters. At the second session of the forthcoming St. Louis meeting of the Society, an opportunity will be afforded, however, for exposition and discussion of these social and educational implications.

So far as the history of this Yearbook as an official undertaking of the Society is concerned, the topic was first proposed at our New Orleans meeting in February, 1937, when Dean Stoddard was appointed chairman and given a preliminary grant of $600. By August of that year a Committee had been named, which, with two or three changes, became the official Committee of the Society. In 1938 there was granted $500, and in 1939, $400. The Chairman has held the Committee well to its work; the schedule has been followed punctually; the material reached the editor in uncommonly good form. The members of the Committee have shown a fine professional spirit; without disparaging any other members I think it not improper to say that Dr. Hollingworth must have devoted almost the last days of her life to perfecting her contributions to this volume, and that Dr. Terman made a, to him, arduous trip to a special Committee meeting at Chicago in order to secure a careful presentation of points of view that seemed to him cardinal.

This leads me to a final word on the contents of this *Thirty-Ninth Yearbook*. The treatment will seem ultra-technical to many readers not professionally concerned with the nature and investigation of intelligence. It is technical. The editor, at any rate, felt that not infrequently he had walked off the edge and was likely to become *spurlos versunken*. There are two justifications that ought to be mentioned; one is that it is the proper function of this Society to devote an occasional yearbook to the more abstruse phases of educational thinking; another is that it is by no means unprofitable for the practical worker to peruse pages that he does not always understand, if only he reaps thereby the conviction that the fundamental and underlying truths about human nature are not so simple after all, either in their formulation or in their discovery.

G. M. Whipple, *Editor*

INTRODUCTION

INTRODUCTION

George D. Stoddard
Chairman of the Committee
Director of the Iowa Child Welfare Research Station
State University of Iowa
Iowa City, Iowa

I. Organization and Purpose of the Yearbook

1. General Organization

The purpose of this introduction is simply to aid the orientation of the reader, not to evaluate the main findings or conclusions.

In general, expository matter by Committee members and their associates has been placed in Part I, and Part II has been reserved primarily for researches, presented either in detail or in fairly substantial summaries of current materials.

2. Some Differences between the Society's Two Yearbooks on Intelligence

The original purpose of this Yearbook was to throw light upon *the nature and nurture of intelligence.* It was felt that the phrasing of this title did not beg the question of demonstrable influence upon the course of mental growth through environmental impact, inasmuch as the term 'nurture' may be reduced to a negligible factor, or extended to include negative effects.

As a bridge between the two enterprises — the 1928 and the 1940 Yearbooks — a citation from Dr. Terman's introduction to the *Twenty-Seventh Yearbook* of the Society may be appropriate:

> Possibly one misconception (now fortunately clarified by recent contributors to the nature-nurture field) has clouded the pertinent issues as much as any single other. This is the notion that a 'nature' theory implies a type of 'glass bottle' mental development which is quite independent of any stimulation from environment. No idea could be more misleading, and no phenomenon more impossible. The theoretical as well as the practical nature-nurture problem that faces us deals with human beings as they are — not as an experimenter might rear them in a nurture-proof laboratory. We are interested in the child or adult as he comes to us

> — with his unique complex of ancestry, associates, home training, schooling, and physical and moral attributes. We are interested, not in finding out how he would have developed if he had had no environment at all; rather we wish to discover whether or not he can be made a more intelligent individual or a more learned one by improving the condition of his *milieu* within the limits found in reasonably good social communities. More generally, we wish to find the relative potency of all types of human environment to add to, or to detract from, human endowment, and to know the limits placed upon achievement by endowment.[1]

As in 1928, so in 1940, with perhaps one fundamental difference in the formulation of the problem. In the 1928 Yearbook, as in most studies on heredity and environment, an answer was sought to the question: What are the relative weights of the factors, nature and nurture, in the mental growth of the child? This question led to statistical approaches based largely upon correlational technique and the analysis of variability. Studies of twins afforded the classic examples.

Increasingly it seems important to ask this question: Regardless of the proportionate weights of nature and nurture (something never to be satisfactorily answered), what are the limits of environmental impact in producing demonstrable effects upon mental growth? Following this lead, research-workers are basing their programs upon attempts to give the young child favorable growing facilities.

3. Contributions Necessarily of Unequal Significance

Very likely no member of the Committee feels that justice has been done to either the theoretical or the experimental aspects of this great problem. In spite of a Committee search for contributors who had approached this subject on a long-time and systematic basis, it is apparent that the contents of the Yearbook are somewhat opportunistic. They do not all bear with equal directness upon the subject of the Yearbook, nor are they of equal weight with respect to validity, clearness, or fruitfulness of outcome.

The reader may note an extreme condensation in certain portions of the Yearbook along with detailed accounts, in standard dissertational language, of certain other (chiefly original) studies. This irregularity in allotted space is due to the conviction of the Committee that it would be well not to reproduce *in extenso* materials that had been published

[1] "Nature and Nurture: Their Influence upon Intelligence." *Twenty-Seventh Yearbook, Part I,* of this Society, 1928, pp. 2–3.

in detail elsewhere. For studies thus abridged here the reader can find a full account in the original publications cited in the bibliographies.

4. Nature of Criticisms

Criticism will be seen to revolve primarily about research material previously published. In this sense the Yearbook is not so much self-critical as pre-yearbook critical. To a slight extent this tendency has been corrected by the inclusion of critical summaries prepared by Committee members.

The reader may feel a lack of balance in the critical portions of the Yearbook. Some original manuscripts submitted late have not been criticized at all by the Committee members. This does not mean that these studies are considered models of research procedure, or that the Committee members concur in the conclusions or interpretations of the respective authors. But all manuscripts were endorsed by at least one Committee member, and on this principle, developed in the Committee and approved by the Secretary of the Society, there have been no exclusions of manuscript.

5. No Attempt at Consensus within the Committee

In the early stages of the Committee's work it was hoped that the Committee itself could arrive at a consensus on scientific outcomes. It is true that, in spite of the divergent findings and views reported in the various studies and expositions, there are certain areas of agreement — perhaps the chapters by Dr. Freeman and Dr. Carmichael are illustrative. The reader may determine for himself what these areas are and how much agreement was revealed. It was not feasible, however, as a group, to attempt the formulation of an explicit and substantial document embodying a composite view. Lacking this, it has not seemed helpful to list, in colorless fashion, the usual generalizations on intelligence and mental testing to which any group of psychologists would give adherence.

What might be called a 'democratic' method, then, has been followed in the construction of the Yearbook. Almost all new materials available have gone in; almost nothing has been kept out. The advantages, as well as the disadvantages, of this procedure are probably apparent to all. The reader obtains, at least, a panoramic view of what is going on, of the issues that seem important; he does not find emergent any array of fundamental conclusions, any authentic single document that all Committee members may endorse, or vary from, in clear and systematic fashion.

To the reader, in short, is left much labor. Perhaps this is no great misfortune.

The Committee members, associates, and contributors who are responsible for the Yearbook have been listed. Acknowledgment should also be made to Miss Dorothy E. Bradbury, Research Associate in Publications in the Iowa Child Welfare Research Station, who assumed the considerable burden of assembling the manuscripts, coordinating the bibliography, and preparing the final printer's copy.

II. Some Preliminary Comments upon the Nature and Nurture of the IQ

The best IQ, like the best medical diagnosis, should be the one obtained nearest the time of a practical decision. From the standpoint of utility, the greatest demand comes for valid testing around ages two (talking), five or six (school entrance), twelve (junior high school), and sixteen (compulsory school limit).

Very likely the importance of the IQ at any age level has been overrated. After school entrance it is clear that mental testing adds little to that knowledge of child ability that could be secured by an analysis of academic records *based on reliable and cumulative measurements in standard school subject matter*. The high overlap between mental ability and scholastic achievement — when the conditions of teaching and examining are good — has led to the abandonment of the accomplishment quotient.

After all, intelligence tests are to a considerable degree simply knowledge tests scored relatively to the achievement of like-age children. Revisions of the Binet-Simon Scale are essentially indirect measures of scholastic aptitude.

Native intelligence, like *native running ability,* is a postulate; the counters actually employed in mental testing are such school and preschool familiars as bead stringing, block building, puzzles, animal pictures, conundrums, counting, definitions, and reciting from memory. The most important single measure in the Stanford Revision is the vocabulary test.

All the 'standardization' and all the involved statistical superstructure erected upon intelligence testing should not erase this fact from our mind: speeding up or slowing down in a child's ability to perform the tasks listed above *is a change in IQ*. Any such change in rate causally related to environmental impacts, of whatever kind from conception to maturity, is all that is meant by *nurture.*

Nor are these tasks mean or superficial; they are important in the child's school-centered world. Nevertheless reëxamination of the contents of tests is a wholesome exercise for those who grow lyrical over the human advantages to be derived from IQ raising, without benefit of mental hygiene or social reconstruction; and a soothing experience, too, for those who grow angry defending the citadel of IQ inviolability.

It is possible that the popular interest in euthenics has resulted from a knowledge of the failure of programs for human betterment through the utilization of eugenical, biological measures: the 'best people' are not having the largest number of children (but the smallest), and human mating is proceeding in complete ignorance of the gene components of the parties involved. When we speak of *good breeding* in horses, however, we do mean just that sort of control: the close inbreeding of selected strains over many generations, with preservation of the fit and a ruthless destruction of the unfit, in terms primarily of physical traits. When we speak of *good breeding* among human beings, we refer really to *good rearing;* the kind of breeding necessary to establish and maintain pure-bred human stock would be considered incestuous.

In any case, democracy is upsetting to genealogy. What can one say about a socio-biological system that, upon a basis of n generations of individual obscurity, produces in $n + 1$ or $n + 2$ generations notable reserves of human talent, leading to the highest achievement along scientific, artistic, and social lines? Certainly, if a political system (democracy) can unleash such forces in a population biologically mongrel in type (the human race), our efforts may center less in the putative quality of human materials and more in the opportunities presented for stimulation and development. The great social problem of the world today is not shortage of talent, but wastage of talent.

SECTION I

DEFINITION OF CONCEPTS

CHAPTER I

THE MEANING OF INTELLIGENCE

Frank N. Freeman
Dean of the School of Education
University of California
Berkeley, California

I. The Concepts, or Definitions, of Intelligence

The clarification of the concept of intelligence and the agreement upon a definition, or definitions, are fundamental for the purposes of the Yearbook. A question of semantics is involved. In the discussion of any problem it is essential that the concepts designated by the terms used be constant. The same person must mean the same thing by a term in different parts of his discourse and different persons must mean the same thing by a given term. In the present case failure to define the term 'intelligence' might result in the assumption, either consciously or unconsciously, of a variety of meanings and a consequent confusion of thought and a prolonged debate upon false issues. Again, it would be possible to set up a definition that would determine the issue and would shut out the consideration of pertinent facts. It is essential, therefore, that, first, the term be defined, and second, that it be defined in such a way as to permit the consideration of evidence pertinent to any important problem concerning intelligence.

I believe we may distinguish broadly between three concepts of intelligence — excluding merely popular concepts or those entirely irrelevant to our problem. These three concepts may be called the 'organic' concept, the 'social' concept, and the 'psychological' or 'behavioristic' concept. The psychological concept in turn includes a number of varieties. These three types of concepts are not necessarily to be judged as either right or wrong. Each of them, I believe, is useful for certain purposes. Each of them may serve as a basis for scientific investigation and each may serve as a term in scientific generalization. Confusion results, however, when any of them is substituted for one of the others in a case in which the other is the appropriate concept. All three, each in its appropriate setting, are probably necessary to a complete understanding of human behavior. It would be better to use a different term for each of the three concepts, but for the present the term is used in the three senses and it is necessary to describe these three.

1. The Organic Concept

As an organic concept, intelligence means the potentiality of a given type of behavior inhering in the bodily constitution of the individual. The kind of behavior made possible by the organic constitution or structure that goes by the name of 'organic intelligence' needs, of course, to be designated. Organic potentiality for flying, for example, or for walking would not be called intelligence. For the moment we may merely say that the type of behavior that we have in mind when we are concerned with intelligence as an organic concept is the same type of behavior that we have in mind when we are concerned with intelligence as a psychological or behavioristic concept. The organic definition assumes that there are differences in organic constitution, and it may assume as a corollary that these differences are hereditary, though this is not absolutely essential. The absence of any constitutional differences that would affect behavior would nullify this concept or at least would represent the limiting case in which inherent individual differences were zero. Practically speaking, it assumes that such differences exist.

This organic concept does not assume that behavior we call 'intelligent' is determined solely by organic constitution. It is compatible with the view that behavior is determined jointly by the constitution and the environment. When the term is used in this sense, however, it refers to the constitutional component rather than to the environmental component or to both jointly.

2. The Social Concept

The social concept of intelligence refers to a certain group of factors that are the outgrowth of social intercourse and social organization. They are sometimes called 'social institutions,' and together make up what is called 'culture.' As in the case of organic intelligence, not all social institutions or all elements of culture are included in social intelligence. Only those leading to behavior that is defined as intelligent in the behavioristic or psychological sense are included. Thus it may be the practice in one community to shave the head and in another to let the hair grow long, but neither practice may contribute to social intelligence. Many elements of culture, however, have an obvious bearing on the behavior we call intelligent. They include sets of symbols, such as language and number; concepts, such as matter, space, law, or duty; and generalizations, such as those expressed in the law of the conservation of energy or the law of gravitation. The individual's effective intelligence — that is, his ability to deal effectively with the problems confronting him — is affected enormously by such social factors as these. In the sense in which the term is here used, the member of a highly cultured or civilized community is much more intelligent than is any member of a primitive community, although in the organic sense he may be less intelligent.

These two concepts, the organic and the social, are not contradictory, but they are different. Great confusion has often resulted from the fact that persons debating the issue of intelligence, its meaning and its factors, have employed the term in these two diverse senses. Psychologists, for example, have sought to demonstrate that differences in intelligent behavior may reasonably, or perhaps even certainly, be ascribed to constitutional or organic differences, sometimes to inherited constitutional differences. Other psychologists or sociologists have sought to prove that similar differences in behavior may be ascribed to differences in culture. Both efforts are perfectly legitimate forms of scientific inquiry, provided the evidence submitted meets the canons of scientific research. Furthermore, the two efforts are entirely compatible with each other. There is no logical reason why individual differences in intelligent behavior may not be found to be due both to organic and to cultural factors. And yet scientific writers and commentators frequently assume without the slightest ground in either fact or logic that to demonstrate the existence of one factor disproves the existence of the other. In some cases the still more devious course is pursued of appropriating the term 'intelligence' to designate the assumed factor alone and refusing to recognize that it may also refer to the other. This is the error referred to in the first paragraph.

3. The Psychological, or Behavioristic, Concept

Psychologists have commonly used the term 'intelligence' in a third sense. They take it to mean particularly the type of behavior that is defined by a given type of test. It is assumed that test behavior is also symptomatic of behavior in life situations. This relation between test behavior and life-situation behavior is a matter of investigation, but behavior is taken and analyzed as it is found without attempting to go back of behavior and to determine its factors. Those who decline any attempt to define intelligence sometimes really assume this definition. Intelligence is really defined, but it is not defined in terms of its underlying factors; rather, in terms only of its manifestations. This is entirely acceptable provided it is not implied that the examination of the concept in either its organic or its social sense is not a worthy subject of investigation.

It is evident that the organic and the social concepts are of a different order than is the psychological or behavioristic concept. The first two designate factors in behavior; the third designates a type of behavior. The first two are explanatory, the third is descriptive. The first two refer to supposed causes of variations in intelligence; the third refers to the form of behavior in which variation occurs. The first two are properly not definitions of intelligence at all, but rather hypothetical factors. We therefore propose that intelligence be defined in the behavioristic sense and that so-called 'organic' and 'social' intelligence be spoken of as *factors* in intelligence rather than as intelligence itself. The purpose of the Yearbook is to explain the nature and the nurture of that type of behavior which we call intelligent. This will lead us to explore the bearing on intelligent behavior of organic constitution and of culture and social institutions, but we shall mean by differences in intelligence, unless otherwise specified, differences in behavior rather than differences in these factors in behavior.

We have indicated the general nature of the concept of intelligence by saying that it is a characteristic of behavior. We have, however, not yet indicated what kind of characteristic it is or what type of behavior it refers to. We have used the term 'intelligent behavior' and have thus begged the question unless we can go farther and say what type of behavior we call intelligent.

a. Two Points of View with Respect to Intelligent Behavior. We may now distinguish two points of view from which behavior has been called intelligent; namely, the *analysis of the behavior itself,* and the *description of its functional relation to the environment.* As before, these points of view may not be contradictory, but confusion results if they are used interchangeably without indication of which point of view is adopted.

Much of the confusion that exists regarding the meaning of intelligence arises from the mixing of the functional and the analytical modes of approach. The functional point of view considers the relation of behavior to the attainment of some end that meets a need or satisfies a desire. The only criterion that enables one to judge whether or not behavior is intelligent is this one of use or value. It may in some cases be difficult to judge whether behavior is useful or to compare the usefulness of different forms of behavior, but to judge behavior or to say whether it is intelligent without the criterion of usefulness is impossible.

b. The Analytical Approach and Its Relation to the Functional Approach. The analytical approach seeks to find out the elements or components of ability. These components are commonly called 'factors' and the process by which they are identified is called 'factor analysis,' but the term 'component' is to be preferred because the purpose of analysis is to discover the parts of which the complex of abilities is comprised rather than to get at factors.

To illustrate: In the functional approach we may inquire whether a given form of behavior, such as perceiving likenesses or differences, using language or numbers, or solving puzzles, is a means to gaining a useful end, such as securing food, escaping enemies, earning a living, gaining prestige, or attaining satisfaction in some other way. In the analytical approach one examines the multifarious types of behavior that are used to gain various ends and seeks to classify them into a few relatively homogeneous forms. This may be done by observation or by elaborate statistical methods. Thus one may conclude that all functionally useful behavior may be accounted for by combinations of such elements as memorizing, perceiving spaces, performing number operations, using language in two or more ways, making deductions, making inductions, and so forth. These components are discovered by noting that differences in effective performance on the part of different persons may be described in terms of differences in these operations or combinations of them.

It should be clearly recognized that analysis is a step in the complete functional description of behavior. It can never be carried on independently of a functional examination. When analysis appears to be independent, it is because the functional aspect of behavior is assumed but has not been explicitly stated or examined. In the usual procedure it is tacitly assumed for the purpose of making the analysis, and the explicit examination of the functional significance is postponed until after the analysis is completed. It may be questioned whether this procedure does not distort the analysis or make it incomplete, but this is not the place to press this question.

The existence of the assumption of the functional point of view in analysis is revealed by the use of the term 'abilities.' An ability implies an end. An ability means nothing except as it is an ability to do something. A higher ability means doing the thing in a way better suited to attain an end than in some other way. Thus ability to shoot implies that it serves some purpose better to hit a target than to miss it, and that performance may be graded in terms of the accuracy with which the target is hit. Similarly with the ability to run a maze, to write, spell, add, typewrite, learn history, or make a speech. Unless it is assumed that performance as measured by a scale of scores or ratings in one direction represents degrees of usefulness, the concept of ability is meaningless.

The actual practice of analysis reveals an additional assumption, or at least raises a question. The abilities that are actually analyzed are abilities to pass tests. Performance of practical activities is rarely, if ever, used. What is really analyzed is test ability. It is assumed that test ability is useful. This assumption is grounded on previous experience with the tests, consisting of comparisons between test performance and other activities, usually scholastic activities. Further validation of the value of test abilities is put off until they have been analyzed and new tests devised to measure the simplified abilities. Such validation is necessary to establish the functional significance of the analysis.

Analysis, therefore, does not contribute to the definition of intelligence. It gives us information as to the composition of abilities, but what types of behavior are functionally useful and what types of functional usefulness should be designated as 'intelligent' are to be judged on other grounds. To evaluate the functional usefulness of behavior, it is convenient to divide it into types or classes, and an improvement of the classification clarifies the evaluation, but actual evaluation requires a study of the relation of behavior to its ends or purposes.

c. The Functional Approach: Adaptation. The most fundamental concept by which to judge the functional significance of behavior is probably that of adaptation. An intelligent act is one that is adapted to its purpose, one that accomplishes its purpose well and economically. An intelligent individual is one who adapts himself to the conditions under which he lives. Such adaptation is necessary for physical survival and well-being and also for the attainment of such purposes or the satisfaction of such needs and desires as the individual may have. Adaptation does not mean the determination of aims or purposes by the environment. These may be set by the organism or by the organism jointly with the environment. Their origin does not matter. The point is that intelligence assumes the existence of aims, purposes, desires, or needs, and the intelligence of a given type of behavior is measured by the effectiveness with which it satisfies them.

The concept of ' adaptation ' does not imply subservience or time-serving. It means merely that the world presents stubborn facts that cannot be ignored but that may be used. It is the part of intelligence to use them.

d. Limitation of Intelligent Adaptation. This makes of intelligence a broad concept; indeed, unless it is qualified, it is as broad as the range of adaptation. This is too broad, since it includes some forms of adaptation that are purely physical. For example, the body becomes adapted to changes in temperature, in food, and in types of work or activity, but this we do not call intelligence. Such adaptation is produced by parts or organs of the body acting somewhat independently of each other. The adaptation we call ' intelligent ' is made by activities that more commonly engage the body as a whole and that consist of definite, observable movements. (In the case of mental adaptations the movements are projected into the future.) These measurements employ the skeletal or so-called ' voluntary ' muscles and are controlled by the cerebro-spinal tract of the nervous system.

e. Routine and Novel Behavior: A Further Limitation. Left in this form, however, the definition is still too broad. We must make another distinction; namely, that between routine behavior and behavior which is to some degree novel. Unless some degree of learning, invention, or improvement in behavior is involved, we do not call it intelligent. By convention it has become customary not to describe the relatively unlearned and relatively invariable behavior of an organism as intelligent. Thus it is advantageous for the human organism that the pupillary opening of the eye should alter in size in relation to the amount of light falling on the retina, but this iris-light reflex is not commonly considered an intelligent act. An intelligent act in general is not an act that one is equipped to perform innately or by virtue of the mere growth of the organism; it is one that the individual has to learn to perform, and one that, when learned, enables him better to meet his needs and satisfy his desires.

This definition would identify intelligence as *the ability to learn, or to perform acts new to the organism.* This is really implied in the idea of adaptation. To adapt to the environment implies that the individual makes some change in his accustomed activity or does something new or different. His action may be adapted (that is, suited to the environment by inheritance or previous learning), but we commonly take 'intelligence' to mean the process of adapting, not the state of being adapted.

f. The Resultant Definition of Intelligence as a Psychological Concept. Intelligence, then, is the ability to learn acts or to perform new acts that are functionally useful. This definition leads to a distinction between types or forms of intelligence to fit the diversity of kinds of functionally useful acts. Different kinds of acts may be functionally useful to animals and man, to man at different stages of development, or to man in different situations. Some would object to making the concept as broad as this. They would confine it to what we ordinarily call the intellectual; that is, to abstract thinking. This, however, seems to be an arbitrary restriction of the term. There is a difference in kind between organic adaptation and behavioral adaptation, but there seems to be only a difference in degree between the different types of adaptation that involve learning. We would include, then, such diverse types of learning as are involved in manipulation, performing an act of skill, identifying an object, learning names of objects, forming concepts, and solving puzzles or problems of all sorts. These all are evidently means of functional adaptation.

It is, of course, possible, and indeed probable, that these forms of learning and of adaptation form a hierarchy. The hierarchy may be of a genetic order and also of a functional order. That is, some may come later in evolution and in the development of the individual, and some may be superior in so far as they enable organisms that possess them to control those that do not. Doubtless there is such a hierarchy. This, however, does not imply that the higher forms can take the place of the lower forms or that a complete measure of intelligence can consist only of those tests that measure the higher forms. The concept of intelligence and tests of intelligence should be broad enough to include the entire range of the hierarchy. Different types of measures of intelligence may be used to measure the ability to make various types of adaptation.

II. The Construction of Measures of Intelligence

If we set out to make a complete set of measures of intelligence we would first try to discover the various kinds of functional adaptation the human being is required to make and would then seek means to test the ability to make these adaptations. The task has not been approached in this systematic way. The form of adaptation that has been most thoroughly explored is success in school. This was first used as a criterion for the validation of tests by Binet. It has since been extensively used for this purpose by many investigators. In the meantime success in vocational pursuits has also been used and there has been some attempt to measure 'social intelligence.' All this is pertinent to the study and the measurement of intelligence.

We may, of course, proceed in the opposite direction; that is, we may develop tests of all sorts and thus secure measures of functional adaptation to artificial test situations. Afterward we may see how far the tests measure ability to adapt to life situations. This is the method that has usually been followed. It is now being pursued in the form of factor analysis. If factor analysis is used as a means of getting a complete measure of intelligence, or functional adaptation, it assumes that test situations can be set up that will require the entire range of adaptations required in life situations.

Making this assumption, factor analysis is an analytical approach that consists, in the first instance, in an internal analysis only. It is concerned only with a comparison of different types of performance, particularly of test performance. It strives to determine resemblances between different performances and thus to discover factors in test performance. The aim is ultimately to reduce all performances to a limited number of factors. These factors may be more or less special or more or less general, and there may or may not be a general factor running through all the performances.

The ultimate meaning of this analysis is to be found in the determination of the functional relation of the different factors, but this determination may come after the internal analysis. In the meantime, the identification of the tests that represent different factors is accompanied by a provisional description of the abilities believed to be represented in these tests.

Studies in factor analysis have thus far failed to produce a unanimity of opinion. Three broad types of concept may be distinguished. In the first place, general intelligence or general intellectual competence is thought of as a composite of a large number of highly particularized abilities. This is the view of Thomson and of Thorndike. In the second place, ability may be thought of as composed of a limited number of primary abilities, each one of which is relatively independent of the others. No one of these may be called intelligence, but intelligence may be thought of as a sum of all, since no one of the abilities is general. This appears to be the view of Kelley and of Thurstone. However, some of the recent works of Thurstone's students suggest that one factor may be more general than the others. The third concept is that ability is made up of a large number of very specialized factors, a limited number of group factors that run through a number of operations, and one factor that is general. This is the view of Spearman and of Holzinger. Whether the general factor should be identified with general intelligence is a question. The general factor is, however, described by Spearman as including the eduction of relations and the eduction of correlates.

III. The Meaning of Intelligence for the Purposes of the Yearbook

For purposes of this Yearbook it is important that the concept of intelligence be kept broad and that it be kept functional. The task of the Yearbook is to discover the factors in intelligence, thus conceived. Its purpose is to find out to what extent and by what ways human beings may be enabled to make better functional adaptation; that is, to serve better their needs, aims, and purposes. It must study the differences between persons in their ability to make this functional adaptation and must seek to explain why these differences exist. We must inquire whether and how far these differences are organic and how far they are social in origin. In explaining these differences light will be shed on the question of the extent and limits of improvability in the individual and in the group.

In seeking answers to these problems we must use the measures of intelligence now available. If these measures are not complete, we must evaluate them in terms of a comprehensive concept of intelligence. We must choose the tests to be used and interpret their results in the light of such a comprehensive concept. The fact that some tests have been called 'intelligence tests' need not restrict our use of them, and we should not let the label that has been placed on the tests decide whether they serve our purpose. Neither should we allow our thinking to be restricted to one of the theories as to what particular mental process or processes constitute intelligence from the analytical point of view. Any mental process that promotes functional adaptation must be accepted. It is the nature and nurture of intelligence in this broad sense with which we are concerned.

CHAPTER II

THE MEANING OF ENVIRONMENT

BETH L. WELLMAN
Professor of Child Psychology
Iowa Child Welfare Research Station
State University of Iowa
Iowa City, Iowa

In studies of the intelligence of children the term 'environment' has been used with a variety of different meanings. In some studies the authors recognize that the use of the term has reference to that portion of the total effective environment under investigation at that particular time in that particular study. Thus, gross changes or differences in conditions or in selected details of the environment are described and related to changes or differences in intelligence. This procedure can be followed with the express recognition that other environmental factors than the ones under study may be producing simultaneous effects upon the intellectual development of the children included. By various means the authors can then proceed to determine whether, and to what extent, the changes or lack of changes in intelligence can be traced to the environmental variables under special consideration.

In other studies, however, when an attempt is made to determine the relative contributions of heredity and environment to child IQ, it has been necessary to deal with all environmental factors considered as a whole. The procedures involved in such cases will be discussed in this chapter under the heading of the relative contributions of heredity and environment.

It can readily be seen that the description and measurement of all of the possible environmental forces operating upon the child during any given period in his development entail a task of Herculean proportions. Even more difficult are the description and measurement of all the accumulated forces that have been operating upon him at various times in his past. Yet the end product, or net result, in terms of child intelligence can be described and measured. Studies have demonstrated that it is possible to cut through the mass of interacting and conflicting trends to at least a rough expression of the net results.

I. Definition of Environment

Webster defines environment as "That which environs; the surrounding conditions, influences, or forces, which influence or modify; Specific.: *Biol.*, the aggregate of all the external conditions and influences affecting the life and development of an organism." Warren's *Dictionary of Psychology* (12) [1] defines environment as "a term covering all physiochemical, biological, and social phenomena which act from without upon organisms." According to Warren, 'environment' is distinguished from 'milieu,' which includes both the organism and its environment. Environmental factors, according to Warren, are "those factors which act upon the organism from without and influence its structure or its behavior."

Philosophers have attempted for years to describe in unambiguous terms what is contributed by the outer world and what by the individual in mental acts such as perception.[2] In general, the systems of thought that may be called 'the idealisms' accept one view of this relationship and those called 'the realisms' accept another. Physiologists and psychologists are now agreed that the receptors, or sense organs, of human beings and infrahuman animals are the main avenues by means of which external environmental factors act upon the total individual, and that a sense organ is only activated by some change in the external energies directly affecting the receptor or by some alteration in the relationship between the receptor and an external energy. These energies are technically described not as 'objects of the real world,' but as vibrations of a given amplitude and frequency per second, and the like. The effective external environment acting upon an organism at any given time may thus be described as the sum-total of the energies that are at that time releasing sense-organ activity.

Besides the term 'external environment' just described, the term 'internal environment' has long been used. This term is not fully satisfactory because it is ambiguous. The internal environment includes those energies inside the outer skin of the organism that act upon the internal sense organs or receptors. Mechanical changes in muscle and chemical changes in food substances in the alimentary canal are examples of such internal environmental energies. The

[1] Such numbers in parentheses refer to the references at the end of the chapter.

[2] The author is indebted to Dr. Leonard Carmichael for the statements contained in this paragraph and the following two paragraphs.

phrase 'internal environment' is also applied to organs adjacent to other internal organs that are at that time under consideration, and especially to the blood stream. Changes in the blood stream are effective in altering the activity of the total organism because such changes affect the sense organs, the nervous system, and the muscles. In this connection, a rapid decrease in the oxygen tension of the blood is described as a change in the internal environment of the brain. Such a change leads to rapid and important alterations in behavior. In certain cases, it is now known that localized regions of the brain are susceptible to slight alterations in the internal environment. It is for this reason that Claude Bernard made his statement that "the fixity of the internal environment is the condition of the free life." It is far from clear, however, that the internal environment in this second sense is to be considered as an initiator of action. It may rather be that it determines changes in conditions that allow external environmental agencies to have a different effectiveness than would be the case under other blood-stream conditions.

In applied psychology, in sociology, in psychiatry, and in everyday speech, however, the world about the organism in its perceptual, or 'real,' form is spoken of as the environment. From this point of view, environment is taken as descriptive of a standard external world of objects that exist in time and space without regard to the location or activity of any special living organism. Thus, a lonesome child may run in terror from a tree stump he observes in the woods because he sees it as a crouching Indian. The energies reaching his sense organs may be the same whether he sees a stump or a man, but the reaction is determined by the way in which he perceives the object. Thus, it appears that 'standard, legal public space and time,' as it has been termed, is determined by the common consent of competent adult human beings. The difficulty with this concept is that it is usually absolute, as for example, 'number of books in the home,' not 'number of books in the home in relation to number of books known by the subject to be in the homes of others,' and the like. In spite of this and other difficulties, it has been found both expedient and valuable to study the relationship between human intelligence and a whole series of environmental factors in this legal, or general, sense. In such studies, however, it must be remembered that there is no guarantee that the environment of the child or of the animal is, for psychological purposes, the same as the environment that is being described and measured. This means that interspecies comparisons and comparisons between individuals of different cultures are rendered especially difficult in quantitative studies of intelligence.

By definition, things or conditions that do not have some kind of effect upon the child are not properly a part of his environment. An example of physical properties that exist but that have no known relation to human activity and development is cosmic rays. The child must be in some active relation to the factor or condition. Environment thus does not exist in the abstract apart from the child. This does not mean that the child is aware of the factors or of his own reaction to them. For example, monoxide gas has a deadly effect upon the organism, of which the individual concerned may be unaware.

Subdivisions of total environment may be made according to those aspects of a child's development that are most affected by a given sector of his environment. From the standpoint of this Yearbook, there might be two major subdivisions: (1) sectors of environment that are nonintellectual in the character of their effects, and (2) sectors of the environment that are intellectual in the character of their effects. The term 'environment' has no meaning in reference to intelligence unless there are intellectual effects. On the other hand, everything that has an effect upon the development of intelligence is environmental unless it is carried in the genes. It is difficult to make a clear line of demarcation between the two subdivisions or to conceive of any sectors that are completely nonintellectual in effects. Certain conditions or factors, however, are preponderately related to physical, rather than to mental, functioning. For example, temperature variations within the ordinary ranges of comfort are probably slightly, if at all, related to intellectual endeavor, although they have certain effects upon the organism. Differences in the color of walls and the presence of beautifully bound books that are never removed from the shelves are other examples.

Instead of pressing the distinction too far, it is probably better to express environmental factors, or sectors, in terms of degrees of their suitability, appropriateness, relevance, or effectiveness in intellectual areas. For example, there is probably a wide range in the effectiveness of books as stimulating factors in children's intellectual development. Minimal effectiveness may be achieved in a home where there are some books, but not books suitable for the child, as, for instance, only difficult adults' books. Maximal effectiveness may be achieved from books accessible for the child's use, and appropriate in their contents, especially if he receives constructive guidance in reading.

Thus, not only should an environmental factor be expressed in terms of whether or not it has in general an effect upon intellectual growth, but also in terms of how much effectiveness it has. The direction of effect may be stimulating, thus raising the IQ; or depressing, thus lowering the IQ. Further, the child is a growing, constantly changing individual. The same environmental factor may, therefore, in reality be quite different psychologically as it impinges upon him at different stages in his development. The same books in the home are not the same environment for the child at age four and at age fourteen.

A chief problem of this Yearbook, therefore, is to locate and describe the factors or conditions that have relevance to mental development and to express the degree of mental stimulation, if any, of different constellations of environmental factors.

Investigations should be pointed in the direction of a better understanding of the dynamics of environmental forces and toward the development of measuring instruments more suitable for application to individuals. For example, it is important to know what actual differences in environment exist between average and very superior children whose fathers have the same occupation.

II. Types of Environment Studied in Relation to Intelligence

Evidence regarding the effects of environment upon child intelligence is presented in Chapter VIII. Here the major environmental constellations that have been treated in the literature are mentioned briefly, with the emphasis on the concept of environment that is involved. The adequacy of a measure of environment is to some extent gauged by its sensitivity in recording effects. That is to say, one criterion of whether variables in an environment are real and properly differentiated is that differential effects are produced by variations in such an environment. *Superior environments, like superior diets, are those environments that produce superior results.* On the other hand, lack of noticeable effect may simply mean that the particular variables as measured are not important from the standpoint of intellectual growth.

1. Socio-Economic Status

Socio-economic status is one of the constellations of environmental factors that has been perhaps more widely studied than any other. Most often this status is expressed as occupation of father, although occasionally as income, wealth, or cultural status of the home. It has been found consistently that groups of children whose fathers are in the upper occupational levels have a higher mean IQ than groups of children whose fathers are in the lower occupational levels.

The interpretation of this relation between socio-economic status and child IQ, however, has never been clear, since better heredity and better environment tend to go hand in hand. It would be unwise to attribute the usual relationship for children living with their own parents to environmental conditions alone or to heredity alone. Further, the interpretation of differences in IQ of children within a given occupation is not clear, since the explanation could be either (1) that there is considerable range in heredity within each occupation, or (2) that there is considerable range in environment within each occupation, or (3) that both factors are operative.

Studies of children separated from their parents shed some additional light on the importance of socio-economic status as a purely environmental factor. Practically all the studies of foster children show a higher intelligence of children when separated from their parents and placed in good adoptive foster homes than would be expected from the socio-economic level of their true parents.

2. Education of Parents

Educational level of parents in relation to child IQ has been somewhat less explored than occupational classification. There is, however, evidence indicating a positive relationship here. As in occupational level, the interpretation is not entirely clear, and the relationships are not high enough for individual prediction. Educational level is usually expressed in terms of the grade completed or the number of years of formal schooling had by the mother, the father, or the midparent (average of father's and mother's schooling). Comparisons between different geographical regions of the country and between different decades are complicated by differences in compulsory school attendance, policies of promotion, and the relative proportions that attend the high school from one generation to another. In general, correlations with mother's education tend to run very slightly higher than with father's education, and correlations with midparent education are higher than with either alone.

3. Preschool Attendance

The technique of approach to preschool attendance as an environmental influence has characteristically varied from that used in the preceding categories. In the preceding categories, most of the investigations have taken a cross-sectional view of the relation between an environmental variable and child IQ at a given time in life. In this and the succeeding categories, a longitudinal view is taken; that is, changes from initial to terminal IQ are related to the environmental variables existing during the period under consideration. Comparisons are made of persons experiencing the environmental impacts under study with persons not experiencing them.

Changes in IQ due to nursery-school attendance have been determined by two types of comparison: (1) changes by the same children during periods of attendance and nonattendance, and (2) comparative changes of matched experimental and control groups, in which the experimental children do, and the control groups do not, attend the nursery school. In such a case the term 'preschool (or nursery-school) attendance' is an inclusive characterization of the environment, covering a complex array of experiences and environmental impacts that may conceivably differ markedly from one preschool setting to another. Consequently it would be unwise to generalize from the findings in one nursery school until it is determined whether similar results are obtained in other nursery schools.

4. School and Home

Studies affording direct comparisons of the intellectual growth of equated groups of children attending different elementary schools are rare. What studies there are have indicated that differences exist, but such studies have not demonstrated in what significant respects the school environments actually differ. It is clear that better measures of school environments are imperative.

Studies of foster children have been thought to indicate the influence of home conditions relatively uncomplicated by hereditary factors. In practice, however, when the children have been placed at relatively advanced ages and intelligence tests or comparable evidence have been available at the time of placement, selective placement has commonly entered as a complicating factor; that is, there was made a definite effort to fit the child to the home, so that the superior children were placed in the superior homes. Only a few studies have appeared in which selective placement was practically inoperative.

It is interesting to note that in the major studies in which children were placed in foster homes in infancy, the mean IQ has usually been found to be high average or superior. The amount of superiority varies in the different studies, as might be expected, but in none of the studies was anything apparent in the case histories or true family background that would justify the expectancy of the degree of superiority found.

5. Geographical Location

Studies of children in isolated mountain communities show the relation between intelligence and environmental conditions by comparing the mean IQ of children in different chronological age groups, and by comparing the IQ with the degree of isolation of the community. The studies report a marked and progressive decrease in IQ with advancing chronological age. Younger children (age 6 and 7 years) are sometimes found to be average in intelligence, while older children (14 to 16 years) are feebleminded. Further, the steepness of the drop is related to the degree of isolation of the community.

6. Institutional Residence

It has been found that children residing in institutions for the feebleminded tend to decrease in mean IQ with increase in chronological age. Until recently an interpretation often put on these facts was that of organic (nervous) deterioration originating within the organism. Recent studies of children transferred from institutions for the feebleminded to orphanages and vice versa give new insight leading to the interpretation of such decreases as a function of a nonstimulating environment, since it was found that individuals so transferred manifested different patterns of change before and after transfer.

7. Training Programs

Training programs set up with direct effort to improve the IQ have been relatively rare. Several studies, however, have reported such programs with high-grade feebleminded and with dull-normal children. Binet (1) described the procedures used and the results obtained by him with children who were three years retarded in school. He was able to reduce the retardation of these children and reported improvement in their general appearance, attitude, and alertness.

In the Binet and subsequent studies reporting favorable effects, certain points of similarity in approach may be detected. Briefly these have been: (1) analyzing the needs and desires of the children, (2) helping them to realize their status and convincing them that they could change their status through their own efforts, (3) progressing educationally from the very easy to the more difficult, (4) adapting the procedures to each child individually, and (5) strengthening personal bonds.

Binet made the point explicitly that no new pedagogy and no new principles of education were involved. Further, he stated that the same approach could be used to good advantage with normal subjects.

III. The Measurement of Environment

Progress in the measurement of environment has been slow. There are several scales available that purport to measure home environment, social status, and socio-economic level with particular reference to school-age children. Among the more comprehensive and widely used measures are those developed by Sims (9), Chapin (4), McCormick (7), and Burdick (2). These have been summarized recently by Leahy (6). Inspection of these scales reveals that they have been largely concerned with the material aspects of environment, and that their chief criterion is possession. The Whittier scale (13) was one of the earlier ones to be developed and has been used by a number of investigators. It includes ratings of homes by the fieldworker on five categories: presence of necessities, neatness, size, parental conditions, and parental supervision. Other scales are similar in general type of content. Leahy (6) included the following divisions in her recent home-status index: children's facilities, economic status, cultural status, sociality, occupational status, and educational status. Correlations between child IQ and such environmental scores usually range between .20 and .50.

Notable exceptions to the tendency to stress material possessions of the home are two scales developed for use with children of preschool age; that of Van Alstyne (11) and that of Skodak (10). Each of these provided for ratings by a fieldworker on the more dynamic parent-child relationships and parent provisions for suitable child experiences.

The first, by Van Alstyne (11), included the following items: suitable play materials and books, conversation of child with adults, proper physical surroundings and routine, other children in the home, association with other children, good economic conditions, suitable excursions, social atmosphere in the home, responsibility for certain personal and household tasks, reading to child, parents' use of good English, educational status of parents, stimulation of independent activity, interest of parents in child's activities, careful response to child's questions, child encouraged to express himself verbally, atmosphere of approval and encouragement, parents tell stories to child, both parents well adjusted and happy in married life, occupation of parents, definite teaching on part of adults, and talking plainly to the child. Van Alstyne found a correlation of .61 ± .05 between the composite of environmental factors and the child's mental age for seventy-five children ranging in IQ from 79 to 167 (mean 113.6). Since this was a group with a restricted chronological age range, all children being between 33 and 39 months, a correlation with IQ's probably would have come close to the figures.

Skodak (10) developed a home-practice inventory for use with foster parents of young children, the scores on which she related to children's intelligence. This scale excluded economic status as such. It included ratings in the following fields of parent practice: play equipment, social contacts, language development, behavior control, parental cultural interests, and social participation. The field of play equipment was rated on the following aspects: creative use of materials, provision of space and equipment for outdoor play, provision of space and equipment for indoor play, quantity of toys provided, stimulation value of toys and materials provided, frequency of parental play with child, frequency of play with other children, success of parties, opportunity for contacts with strangers and new situations outside the home, and stimulation value of trips and other opportunities for expansion of child's interest and curiosity. Other fields included in the scale were similarly developed in detail.

For 103 cases, Skodak found a correlation of .49 ± .06 between the score in parental practices of foster parents and child IQ. She found a progressive increase in mean IQ of the adopted children with each ten-point increase in inventory score. She also found a progressive increase in mean inventory score for each ten-point increase in adopted child IQ.

IV. Relative Contribution of Heredity and Environment

There are two major methods of studying the effects of environment upon child IQ: (1) the study of changes in IQ over a given period under specified environmental conditions, and (2) the analysis of the relative proportions of a child's IQ that are attributable to heredity and to environment. The former is usually longitudinal, the latter cross-sectional.

In the first type of analysis it is necessary only to ascertain what a given set of environmental conditions does to a given group of individuals of known IQ at a given time. The expectation is that heredity will not change in its ongoing effects. 'Inherited brightness' or 'inherited dullness' can be ascertained as well at age four as at age fourteen, provided there are valid tests and samplings. If fluctuation in the rate of growth occurs, heredity may be said to permit this, but scarcely to determine it.

In the second approach it is necessary to measure heredity and also to measure the net result of all environmental forces that have played upon the child up to the time that he is selected for study. Any inaccuracies in one area will throw an undue proportion to the other area.

1. Measures of Heredity

It should be recognized that measures of heredity are not directly based on a knowledge of human genes. Very little is known about human genetics. In particular, no study of heredity in reference to child IQ has presented any evidence whatever on the genetic composition of the particular group under study. Little attempt has been made to go back of the immediate parents, even on the kinds of hereditary measures that have been used. Compared with the elaborate controls often used in animal experimentation, restriction to one generation of ancestors appears inadequate.

The typical psychological study is backward-looking on heredity, beginning with the child and working backward to the parents; the typical animal study is forward-looking, beginning with the ancestors and tracing through generation after generation to the final offspring. The biologist recognizes that not all individuals of a given parentage possess the same characteristics to the same degree. He does not attempt to predict from given progeny the characteristics of the parents.

Lacking actual information on genetic composition, the psychologist relies on resemblances between child and parent. The usual measures employed are not unequivocally hereditary. Parent IQ, for example, could be related to child IQ through environment as well as through heredity. Similarly for occupation of father and extent of education of parents, the most commonly used measures in nature-nurture studies. Recognition of this complexity is given in such studies by the statement that the relationship is a measure of the combined effect of heredity and environment. Comparisons are then undertaken with other groups, such as adopted children, where the hereditary factor is presumably zero, and environment is assumed to afford the sole explanation for such relationships as are demonstrated to exist.

Relationships between child IQ and parental intelligence, occupation, or education are much too low for predictive purposes; the correlations usually hover around .50 and rarely exceed .60. These are the relationships characteristically obtained from the combined effect of heredity and environment. Obviously such relationships leave a considerable margin of IQ variance not accounted for. In the study of identical twins by Newman, Freeman, and Holzinger (8) this problem was frankly admitted. In the Leahy (5) study it was not explicitly recognized.

In the Burks' study an attempt was made to meet the problem on the grounds of logical expectancy from evidence on stature. The final proportion attributed to heredity by Burks departed from the actually measured and statistically obtained figures.

To quote from Burks (3):

> We have now seen that the total contribution of systematic (or measurable) home environment is close to 17 percent, and that the contribution of home environment and parental intelligence together is represented by a multiple correlation coefficient (corrected) of .61, or by a percentage of .37. If not more than 35 or 40 percent of the variance of children's IQ's is accounted for by reference to these factors, what contributes the other 60 or 65 percent? . . . Probably the major share of the residual variance is due to congenital endowment, since in known modes of hereditary transmission the influence of heredity is *always* far stronger than parental correlations alone would indicate. This is necessarily the case because only half the chromosomes of each parent are passed on to the off-spring. Hence, the parental deviation for any trait in question is determined by a number of factors other than the ones transmitted to the child. In hereditary traits, such as stature, which are known to be influenced relatively little by ordinary differences in environment, the multiple correlation of child with parents is .64, but the contribution of heredity to variance approaches 100 percent. The closeness of our estimated value of the 'genetic' multiple correlation for intelligence to this value of the multiple correlation for stature is striking. *Probably, then, close to 75 or 80 percent of IQ variance is due to innate and heritable causes.* (pp. 303–304)

Such arguments by analogy are likely to be misleading. It is not safe to maintain that because the size of r is comparable for two types of characteristics, the hereditary component is equivalent. That stature is 100 percent hereditary is an assumption. If it were true, prediction should increase with the addition of ancestral stature measurements from several generations.

Other factors that might account for a certain percentage of the IQ variance are: (1) differential constitutional factors not carried in the genes, such as prenatal nutrition; (2) environmental factors related to nutrition, endocrine functioning, and metabolism; (3) chance factors undetermined by any known conditions.

2. Inference of Change in Child

Perhaps one of the most serious deficiencies in the nature-nurture studies is the tacit assumption that changes in IQ of the child through environmental influences will be reflected in the correlations between child IQ and other factors to such an extent that the correlations will indicate what the changes have been. Actually, it can be demonstrated that this assumption is questionable. All children in a group could increase or decrease in IQ in constant amounts and the correlation would remain the same. However, the usual assumption appears to be that the changes, if any, will be not constant, but differential in amount. It can be demonstrated that widely different patterns of differential changes can be made without shifting the size of the correlation. Investigators of the nature-nurture problem actually have at hand only the end correlation. From this they must estimate the amount of shift that has taken place.

Correlations done on hypothetical groups with known initial IQ illustrate the difficulties encountered. They emphasize the following points: (1) the size of a correlation between child IQ at the later childhood ages (designated hereafter as the end *r*) and foster-parent IQ depends in part upon whether there was a zero *r* at the time of placement; (2) an initial *r* of zero cannot be assured from random nonselective placement, even by the method of drawing by lot; (3) similar end *r*'s may be obtained from rather widely different distributions of change in IQ; (4) it cannot be predicted from a high end *r* that more children have changed in a large amount or that fewer children have not changed.

In regard to the second point that an initial zero *r* cannot be assured, correlations ranging from .19 to —.14 were obtained in four drawings by lot of hypothetical children. The procedure was to line up a group of sixty-five hypothetical foster parents from highest to lowest IQ. The range in IQ of foster parents was 80 to 140. Slips representing sixty-five children of similar range in IQ were placed in a hat, drawn by lot, and placed in order of the foster parents. The resultant correlations from four trials were .19, .04, —.14, and .05.

Following this, changes in IQ were assigned to the children according to definite principles. The principles of change used here were two: (1) In some calculations changes were scaled according to the amount of difference in IQ between the foster-parent IQ and child IQ, with the provision that the child change in the direction of the foster-parent IQ; (2) in other calculations the changes were scaled according to the initial IQ of the child. When the same principle of change was applied to two groups with different initial r's, the resultant end-correlations (after the change) varied considerably. For example, when the group was used that had an initial r of .04, certain superimposed changes resulted in an r of .56. When the same principle of change was applied to the same children placed so that the initial r was .19, the resultant correlation was .73. Again, when the initial placement afforded an r of .05, the assignment of certain changes resulted in an r of .51, while the same principle of change applied to the same children when their initial r was —.14 resulted in an r of .34. These results demonstrate that the same changes in IQ by the children yield different end r's depending upon the initial r.

Again, it was found that approximately the same end-correlation can be obtained from rather widely different distributions of change. The same children placed identically with the same foster parents were involved in three end-correlations of .57, .51, and .54. Yet the distribution of changes was very different in the three samplings. The initial r was .05. The distribution of changes was as follows.

Change in IQ	*Number of Cases in Which* $r = .54$	$r = .51$	$r = .57$
−10	2	—	—
− 5	3	—	—
0	32	—	2
+ 5	11	12	4
+10	5	12	32
+15	2	15	10
+20	10	12	5
+25	—	14	2
+30	—	—	10
Total	65	65	65

In the first correlation zero change was made by 49 percent of these children; in the second correlation zero change was made by zero percent of these children; and in the third correlation zero change was made by 3 percent. Likewise, the percentages of these children gaining 20 of more points were respectively, 15, 40, and 26 percent in the three calculations. From the end-correlations alone it would be impossible to predict the shifts in IQ.

3. Neglect of Environmental Factors Outside the Home

In nature-nurture studies, as has been pointed out, it is necessary that the net result of all environmental influences that have played upon the child up to the time of measurement be taken into account. Since the children included in the studies have been of school age, it seems reasonable to postulate that school and community influences should not be overlooked. Perhaps the 'purest' measures of home influence will be obtained at those child ages when the children are less subjected to school and community influences; namely, the preschool ages for children who do not attend the nursery school.

Further, it does not seem feasible to group together children from 5 to 14 years of age until it has been clearly demonstrated that environmental influences are not differentiated with respect to child age. It seems reasonable to hold that the same home and the same parents may be more successful stimulators of mental growth at one child age than another.

4. Do Foster Parents and Own Parents Provide Similar Environments?

The point as to whether foster parents and own parents provide similar environments can be raised at the present only as a question. However, there are some grounds for doubting whether similar environments are provided by matched groups of foster parents and own parents, when the matchings are in terms of occupation, education, IQ, and similar measures. There is the possibility (1) that foster parents provide more homogeneous environments, (2) that the fact that they have wanted children and waited a long time for them may make a difference in their behavior, and (3) that the knowledge that they are foster parents may color their behavior and attitude towards the child. Greater homogeneity of environment would tend to reduce correlations between child and parent.

5. The Concept of Maximal Effects

When it is desired to determine the influence of environment on children, it is important that environment be given a maximal opportunity. Otherwise, some of the IQ points rightly attributable to environment will be assigned to heredity. At this juncture the psychologist is in a dilemma, since no one knows what is an optimal environment for a given child or a given group of children. One cannot assume that the particular environmental influences to which a child has been subjected represent the best possible environment for him.

V. How Environment Produces Changes

How does the environment produce changes in intelligence? The answer to this question depends on the nature of the individual child and on the value of the environment for mental stimulation.

1. Nature of the Organism

A first principle in discussing factors governing or limiting the amount and rate of change in intelligence is recognition of the fact that individuals may differ greatly in amenability to change. Individuals must be taken as they are and their individual differences understood. A given child's IQ may remain the same or it may change. However, if the problem of the relationship of the child's IQ to environmental forces has any meaning, it is necessary that some children change more than zero points in IQ.

In order that change may take place, it is necessary to allow some time for the change to occur. Now, since the child is a developing organism that takes a number of years to mature, it is seldom possible to observe and measure him over his total life span, even if there were available adequate measuring instruments for gauging intelligence at every point in this span. Instead, the investigator experimentally sets up boundaries within which to make observations and measurements. The investigator can thus speak with authority only in regard to the particular segment of the child's life and the particular segment of IQ change that his experiment has covered.

A primary consideration in regard to the nature of the organism is initial IQ. This may mean at birth or at any age considered as initial; that is, at the beginning of the period under consideration. The amount of shift in IQ will depend somewhat upon how high it is at the beginning. Theoretically, if the child's IQ is as high as the particular environment is capable of producing in him, there will be no increase; if lower, there will be an increase. Conversely, if the environment is a depressing one, the amount of drop will be greater when his initial status is higher.

In some instances there are structural conditions of the organism that determine limits beyond which it cannot grow. In this category are birth injuries, brain lesions, structural defects of brain or nervous tissue, either present at birth or acquired thereafter, that are serious enough to interfere with mental functioning. Two individuals of 70 IQ might, for example, present entirely different possibilities for future changes, the one having reached his top level, (structural limitations not permitting further increase, although they might permit decrease) ; the other presenting various possibilities for either upward or downward changes depending upon the ingredients of the environment and their adaptation to his particular needs. To these structural limitations may be added the ravaging effects of diseases, like encephalitis and congenital syphilis, that seemingly impair mental functioning.

In addition to initial IQ and physical limitations, there are other psychological differences between individuals that determine the degree of responsiveness to environmental factors. At present we do not know too clearly what these psychological differences are; but they probably include emotional blockings, differences in emotional tensions, variations in goals or levels of aspiration, the individual's feeling of freedom or restriction of movement, and variations in methods or techniques employed to maintain self-respect.

Specific past experiences and the applicability of past experiences to new learning situations may be important, too. If these turn out to be of great importance, then there may be a period in the child's life when he will show the greatest susceptibility to environmental influences. At one stage of his life his psychological make-up may show greater plasticity and fluidity than at other stages of his development. Specific past experiences may conceivably work in either of two opposing directions: (1) they may enhance and give point to a new experience, or (2) they may serve as factors restricting the full utilization of new experiences.

2. The Value of the Environment for Mental Stimulation

Since the child is a changing organism, the suitability of environment for him changes too. This may take place in two ways: either the environmental factors remain exactly the same through a period of time, in which case suitability changes because the child has changed, or the environmental factors change so as to keep pace with the changing child. In the latter case, suitability is judged in terms of whether the rate and character of change of environment are adjusted to the rate of change of the developmental needs of the child.

The environment may change in the same direction as the changing child but at a different rate. If the environment changes at a rate slower than developmental needs, there results a difference in suitability at different ages. For example, environmental factors might have been suitable when the child was two years old, but when he was six years old the lag in environmental change might be definitely noticeable, and when he was ten years old, his environment might be suited to his developmental needs at six years.

Another way of stating the foregoing considerations is to say that the same objectively described factor or condition has different dynamic value at progressive age levels, or that different factors may have the same dynamic value at different age levels. Parental practices must change as the child grows older if they are to be equally stimulating at various ages.

Again, it can be seen that the same objectively described general environment may differ in its suitability for children of different IQ levels. An environment that is geared to dull children will not be a stimulating one for geniuses. This does not mean that the solution to the problem necessarily lies entirely in homogenous grouping on the basis of ability. Merely grouping dull children together does not insure that the environment is entirely suitable for their maximal stimulation; it may, indeed, have the reverse effect. Again, an environment at a level too far in advance of a child may fail to be stimulating.

3. Conditions Imposed by the Investigator

Whenever an investigator sets up a study, he selects children to be studied and environmental conditions to be included, and limits the time of exposure of the children to the environmental conditions. If we think of the total possible 'time of exposure' as a 'time line,' we can see that the investigator necessarily cuts across this line. When cross-sectional studies are made on groups of children of various ages, say five to fourteen years, it can be seen that the potential time line is cut across experimentally at different points for different children. If generalizations are thus drawn concerning environmental effects with such heterogeneous ages, the implication is that the environment has accomplished as much at five years of age as it will be able to accomplish by fourteen years. When the investigator cuts across the time line for both initial and terminal measures, it is obvious that his conclusions on environmental effects must be limited to the intervening period. The initial measurement may have been affected previously by the very environmental forces under consideration, and these same environmental forces may affect subsequent development.

4. Behavior Mechanism by Which Changes Are Effected

In our present state of knowledge any discussion of the behavior mechanisms by which changes in IQ are effected must be in the nature of postulates. However, a few such postulates are indicated by the experimental evidence. They are sketched here as suggestive of the way in which environmental impacts may be translated by the child into appropriate behavior or action leading to IQ increase. The impacts that lead to IQ decrease are the reverse side of the story.

1. First, it may be postulated that children gain in IQ in accordance with the opportunities they find for utilizing the abilities they possess. As has been said, children may differ greatly in their recognition and utilization of the opportunities afforded them. Children who respond to opportunities by full utilization of their present abilities will gain more in IQ than other children who do not put their abilities into effective use in new situations.

2. The second postulate is closely related to the first; namely, that intellectual curiosity is fostered by variety of mental experiences.

3. The third postulate is concerned with motivation. The child's inclination to utilize the resources at hand is of recognized importance. Children generally have a strong desire to maintain a respectable status in their group. It is postulated that the height of the goal of achievement they set for themselves is in part determined by the group or groups with which they associate. Their evaluation of the group and of their own rank in the group determines in part the amount of effort they will put forth towards their own achievement. It is postulated that such effort and such struggle are important factors in mental growth.

Closely allied to the goals set up by the child are the goals or standards held up to the child by parents, teachers, and others. These, to be effective, must, of course, be translated by the child into self-goals. Children generally try to live up to expectations.

These postulates are set forth, not as the final pattern, but as working hypotheses arising from experimentation. They may serve as partial guides for future experimentation.

References

(1) Binet, A. *Les Ideés Modernes sur les Enfants.* (Ernest Flammarion: Paris, 1911) 346 pp.

(2) Burdick, E. M. "A group test of home environment." *Arch. Psychol.*, 16: 1928–1929. 115 pp.

(3) BURKS, B. S. "Nature and Nurture: Their Influence upon Intelligence." *Twenty-Seventh Yearbook, Part I,* of this Society, 1928, 219–316.

(4) CHAPIN, F. S. "Socio-economic status: Some preliminary results of measurement." *Amer. Jour. Sociol.,* 37: 1931, 581–587.

(5) LEAHY, A. M. "Nature-nurture and intelligence." *Genet. Psychol. Monog.,* 17, No. 4: 1935, 235–303.

(6) LEAHY, A. M. *The Measurement of Urban Home Environment.* (University of Minnesota Press: Minneapolis, Minn., 1936) 70 pp.

(7) McCORMICK, M. J. "The measurement of home conditions." *Nat. Catholic School of Social Work, Soc. Serv. Monog.,* No. 1: 1929. 23 pp.

(8) NEWMAN, H. H., FREEMAN, F. N., and HOLZINGER, K. J. *Twins: A Study of Heredity and Environment.* (University of Chicago: Chicago, Ill., 1937) 369 pp.

(9) SIMS, V. M. *The Measurement of Socio-Economic Status.* (Public School Publishing Co.: Bloomington, Ill., 1928) 33 pp.

(10) SKODAK, M. *Children in Foster Homes: A Study of Mental Development.* (University of Iowa, Studies in Child Welfare, Vol. 16, No. 1: Iowa City, 1939) 156 pp.

(11) VAN ALSTYNE, D. *The Environment of Three-Year-Old Children: Factors Related to Intelligence and Vocabulary Tests.* (Teachers College, Columbia University, Contributions to Education, No. 366: New York, 1929) 108 pp.

(12) WARREN, H. C. (Editor). *Dictionary of Psychology.* (Houghton, Mifflin Co.: Boston, 1934) 372 pp.

(13) WILLIAMS, J. H. "Whittier scale for grading home conditions." *Jour. Delinq.,* 1: 1916, 271–286.

SECTION II

DEVIATIONS IN INTELLIGENCE

CHAPTER III

THE SIGNIFICANCE OF DEVIATES

LETA S. HOLLINGWORTH, LEWIS M. TERMAN,
and MELITA ODEN

I

REVIEW OF RESEARCH

LETA S. HOLLINGWORTH [1]
Professor of Education
Teachers College, Columbia University
New York, New York

I. GENERAL CONSIDERATIONS

The scientific literature dealing with intellectual deviates has grown out of all proportion to the numerical frequency of such persons in the population. There are good reasons for this.

As far as education is concerned, it has been verified repeatedly that the great mass of teachers' problems having to do with pupil management centers in the fact of intellectual deviation, more particularly in the fact of mental deficiency. But the problems of the school, in its dealings with the young, are merely premonitory of the perplexities of society, in its dealings with adults. More and more it is realized that 'the mass of men,' those finding their place in the middle 50 to 60 percent of all who are born, create no special problems for themselves, educational, social, economic, moral, or legal. As a group, men of normal (average) intelligence, the 'mass of men,' tend neither to create social problems, nor to solve problems created by the forces of the physical environment. It is the intellectual deviates who create for mankind the great problems of crime, dependency, unemployment, and like difficulties; and who, on the other hand, in the extreme plus degrees of deviation, introduce innovations, evolve civilized life for all (98),[2] and at the same time bring about those inven-

[1] Deceased November 27, 1939.

[2] See the numbered references at the end of the chapter.

tions that endanger and harm mankind by changing the common life in ways for which the mass of men is not prepared or qualified (82).

To establish that these ideas about the significance of intellectual deviation are currently gathering force will not be our main purpose here, and will not be further argued. The fact that such ideas have recently achieved official expression by the National Education Association, through its Committee on Socio-Economic Goals of America (74), and by the National Committee on Coördination in Secondary Education (73) points to the trend of thinking on this matter.

The study of intellectual deviates has also a special interest for psychology as a science, apart from all question of the relation of deviates to socio-economic events. Among extreme deviates, the intellectual processes appear in minimized and in magnified calibers and forms. Thus by the study of processes compared and contrasted at the two extremes of distribution, general principles of cognition may be derived more easily than by other methods. A large part of present knowledge of the principles of cognition and of the nature of intelligence, including the invention of the method of measuring intelligence, has actually been established from the study of deviates.

II. Inferior Deviates

In appraising the *status praesens* of knowledge, inferior deviates will be considered first in order, because they have received first consideration in the scientific study of intellectual deviation. This priority was due to the patent troublesomeness of mentally deficient persons, rather than to an explicit recognition of their primary importance. Indeed, it is highly probable that superior deviates far outrun in socio-economic importance those at the extreme of mental deficiency.

With respect to inferior deviates, we need first of all to know, for practical purposes, what is shown by a census of such individuals.

1. Census of Inferior Deviates

Strictly speaking, all whose intellectual caliber is below the precise middle of a population at large may be called 'inferior.' However, the practice, established about 1916 (88), of defining 'inferior' as a status falling more than —1 P.E. away from the midpoint of the distribution is still, at the date of this writing, accepted in the literature. This automatically includes 25 percent of the school population under the concept of *inferior deviate;* though the concept of feeblemindedness still includes only about 2 percent, those falling at or below about 70 IQ (Stanford-Binet).[1]

It is of perennial interest to know how the incidence of inferior degrees of intelligence is related to factors such as sex, ethnic stock, and socio-economic status. The statistics of inferior deviates herein reviewed are based on censuses taken by the method of mental tests. Such censuses relate mainly to the United States.

2. Frequency as Related to Sex

It has been shown that the preponderance of males in institutions for mental defectives is a function of differential socio-economic pressures, since the males represent a higher selection than do the females (46).

The attempt to compare the sexes as regards frequency at the minus extreme has continued. Bennett (6) found in the special classes of Baltimore about three times as many boys as girls, but in the regular classes the analogous girls were found. Scholastic pressures differ for the two sexes.

In order to study this subject adequately it is necessary to test 100 percent of the persons of each age in a population. This had been approximated in Scotland, where in 1932 all ten-year-olds were tested (85), showing a greater variability for boys. Again, in Scotland all children born on February 1, May 1, August 1, and November 1, 1927, were tested, and in this sampling also the boys showed a slightly greater variability, though the difference was not very reliable.[2] Fraser and others (32) reported in 1935 Otis Tests on boys and girls in Bath, of whom almost perfect samplings were obtained, and found a greater variability for boys.

McNemar and Terman (70), summarizing the literature to 1936, found in " the large amount of data based on more or less nonstandardized psychological tests " 320 cases of greater male variability, 322 cases of greater female variability, and 24 cases of equal variability. Standardized verbal intelligence tests (batteries) gave greater male variability in 29 out of 33 comparisons. McNemar and Terman reach the general conclusion as regards IQ that the standard deviation for girls appears to be 16 points; for boys, 17 points. " This difference of

---

[1] In the remainder of this chapter numerical IQ's will refer to Stanford-Binet IQ's unless some other test basis is designated. — *Editor.*

[2] Scottish Council for Research in Education. *International Examination Inquiry*.. Privately circulated. 1938.

one IQ point in variability would mean that in general about nine boys to six girls would score above 140 or below 60 IQ, and that twice as many boys as girls would exceed 160 or fall below 40." It would also mean that no girls at all would be found at the lowest and the highest points on the curve of distribution of IQ. However, girls are present among vegetative idiots, and a large proportion of the cases reported in the literature describing children testing at and above 180 and 190 IQ have been girls.

3. Frequency as Related to Ethnic Stock

A summary of the literature of Garth (34) indicates that wide differences exist among groups of different ethnic stocks in the population, as regards their respective contributions to the lower ranges of performance on intelligence tests. The same fact is evident in the summary by Anastasi (2), although this commentator dismisses the subject briefly, on the ground that 'racial' classifications are in question, and that selections of emigrants and immigrants cannot constitute an adequate basis for determining what may be 'racially' typical.

The facts that have been derived to date about American school children show that the following distinguishable groups in the United States contribute more than a chance share to inferior degrees of test performance: the American Indian, the Sicilian (Southern Italian), the Negro (especially the Southern Negro), and the Mexican (34).

It must be borne in mind that no survey covering *all* individuals of any of these or of other groups has ever been made. The fact simply is that wherever these groups have been sampled in the United States they show a low central tendency and a very large contribution to low scores on intelligence tests. These results have a bearing only *on the selections of population* received in the United States, and do not inform us concerning the mother populations from which our immigrants came. Whether these group differences are biologically inherent in the ethnic samples received or are functions of differential environment cannot be settled in the present state of our ignorance; crucial studies are lacking, and near-crucial investigations are few.

Bere (7), taking ten-year-olds of three immigrant derivations and matching them with respect to factors often thought to be chiefly influential on test scores, still found significant differences. However, it may be that there were among the groups subtle environmental factors, not taken into account, that may have caused the differences.

Klineberg (60), correlating the mental-test scores of Negro children in New York with the length of time they had been in this supposedly more favorable environment, found that where the families had been longer in the North, the children scored higher. However, in an emigrant movement of this kind very probably the most intelligent come first (*i.e.*, take the initiative); if so, the reasons that those of longest residence have more intelligent children would be genetic instead of environmental, as Klineberg supposes.

The attempts to prove that the superiority of Northern Negro children is not a function of selective migration are unconvincing. The proofs reached were by comparing *school grades* of migrants and of non-migrants, and by testing twelve-year-olds, one group of whom had been in residence in New York two years longer than the other.

The school marks of children clustering at 80 IQ, the median for Southern Negroes, would not be likely to differ materially from those of children clustering at 87 IQ, the figure representing the extreme of the long-residence groups studied.[1] Nor would a difference of twenty-four months be likely to affect appreciably the quality of migrants, where 81 IQ, as compared with 87 IQ, represents the total change claimed as the result of a period of eight years and longer.

In Klineberg's study, as in other studies, the median IQ of the Negro falls into the range below 90 IQ, regardless of length of residence in 'more favoring' environment. The mean IQ for Southern Negroes in various samples studied falls at about 80. Even the Northern Negroes more than eight years in residence, or born in the North, showed a mean IQ of only 87.

If bilingualism, differences in economic ways and means, and other such environmental conditions are adequate to explain why some of our ethnic groups contribute so much more heavily than others to inferior deviation, then it is hard to say why groups like the Japanese do not contribute more than their 'quota.' Strong (89) concluded from a survey of established records that the second-generation Japanese on the Pacific Coast do not differ from the general American stock of that region in mental capacity as shown by tests. If different environments cause the differences found between ethnic groups, why are not the Japanese of the second generation handicapped in comparison with the generality of Americans as much as second-generation Sicilians or Negroes are?

[1] From studies of the inaccuracy of teachers' marks, it seems quite probable that a difference of 20 to 25 points of IQ would be required before a difference in teachers' marks would be effected.

A large area of ignorance lies here, inviting penetration and illumination by scholars of the future. The final establishment of verifiable truth by crucial studies in the field of differences among ethnic stocks as such will be of fundamental importance for general policy in various respects. For the present, American educators have only the fact that those distinguishable and namable groups mentioned above (the American Indian, the Negro, the Southern Italian, and the Mexican, in the United States) yield disproportionately large numbers of deviates who test low on intelligence tests at present available.

4. Frequency as Related to Socio-economic Status

A review of studies, chiefly recent, dealing with socio-economic status and intelligence has been offered by Neff (75). There is general agreement among all such studies that parents who rate low in social status and in economic condition furnish a disproportionately large number of inferior deviates. This fact is not disputed, but the dispute continues concerning the causal relation of the two variables involved.

Neff is strongly biased in the interpretation that the low intelligence scores of the poor and obscure are caused largely by the meager environments in which such persons typically live. He concludes (pp. 754–755):

" All, then, of the twenty-point difference in IQ found to exist between children of the lowest and highest status may be accounted for entirely in environmental terms."

Taking this environmentalistic conclusion as a point of departure, one may note that crucial studies are almost wholly lacking concerning the *cause* of the great number of inferior deviates coming from homes of low socio-economic status. Further, in order to have a complete view of the subject, it would be necessary to supplement Neff's bibliography with references that he neglects, such as those of Teagarden (90), Hollingworth and Cobb (54), and Lamson (63), who did not find that the IQ is raised by improving the scholastic environment, and who failed to equalize achievement in intellectual work by equalizing educational opportunity at the level of the elementary school among children who were from *the same* socio-economic levels of society but who differed greatly in IQ.

Furthermore, the point is missed by Neff that Goodenough (37) found as great a difference between the infants of various socio-economic levels as Haggerty and Nash (40) found at elementary-school ages of such children. If environment explains the differences, then should not these differences *increase* with longer sustained influence of the envirnoment? In fact, Neff uses the *increase* of difference demonstrated by Gordon (38) in the case of younger versus older canal-boat children to prove that environment caused the difference.

For the time being, then, we must recognize that crucial studies of causation in this field are so difficult and expensive that none have been made, and must content ourselves with the mere descriptive statement that poor and illiterate parents, who earn a living at unskilled manual work, contribute a very large portion of offspring to the range below 90 IQ. This fact no one disputes.

5. Physique and Neuromuscular Ability of Inferior Deviates

Recent research continues to verify the earlier findings, which show a low positive correlation between the caliber of intelligence and the magnitudes of physique. The interpretation by Paterson (77), that there is practically no relationship between physique and intellect, does not stand up under the figures from original studies. Within the decade now elapsing, it has been shown by Brander (9) that among prematurely born babies the lower the median birth weight, the lower the IQ subsequently found at ages 7 to 15. Inferior deviates were characterized by a lower birth weight.

In 1930, Davenport and Minogue (20) proposed a 'physical quotient' in which a variety of developmental phenomena would be compared in the individual with a standard set for age. The arithmetical average of status in these several physical traits being found, and being divided by chronological age, gives the 'quotient.' Operating after this fashion, these investigators related IQ to PQ (physical quotient) in seventy-eight feebleminded boys under fourteen years of age and found positive correlations of from .29 to .49 within a range severely restricted for IQ.

Wheeler (102) has presented excellent graphs showing that dull children are slightly inferior, as a group, to the norms established for stature, weight, and several other variables of physique.

Dayton (21) has made an analysis of 14,176 mentally inferior schoolchildren with reference to the incidence of physical defects. The number of physical defects observed increases in this study as intelligence decreases. The same investigator (22) has shown that among mental defectives in an institution the death rate is far higher than for the population at large and is clearly related to the degree of defect; the idiots, for example, die earliest.

Baller (5) followed the subsequent history of special-class pupils in Nebraska, averaging 60 IQ, and found a death rate seven times that of a control group with IQ's from 100 to 120. This indicates that the high mortality rate established by Dayton does not result from institutional conditions, but is a correlate of low intelligence wherever found.

6. Personality Development as Related to Inferior Intelligence

Because the methods of studying personality are still elusive, scientific observation on this phase of the development of inferior intellectual deviates has lagged. The laboratories at Vineland (25, 26, 96) have recently been leading in studies of this kind.

Dating from earlier work by Morrison (72), it has been observed with precision that persons of defective intelligence participate in the emotional experiences and behaviors characteristic of the human species. The work done at Vineland suggests that personality develops in great variety among the feebleminded, though characterized by certain limitations lacking when intelligence is at higher levels. These limitations apply especially to those aspects of personality that are associated with imagination, abstraction, and the appreciation of symbols.

Collmann's work (16) shows nicely how the intellectually subnormal fail, relatively, to be disturbed by symbols.

Here three groups of school children, widely separated as respects IQ, but alike in other essential particulars, were checked with the psychogalvanometer while receiving three various sorts of stimuli. These stimuli were: (1) a warning signal, (2) concrete sensory experiences without verbal accompaniment, and (3) words.

Children below 70 IQ proved to be less disturbed by all three kinds of stimuli than were those from 90 to 110 IQ; but in respect to *words* they were least disturbable of all groups (although the words used were well within their vocabularies).

The comparative stability of the PGR (psychogalvanic reaction) of unintelligent children is speculatively discussed by Collmann. He suggests that the feebleminded children had not sufficient associative power to derive maximal stimulation from the situations offered; while the very bright had such associative wealth that, though stimulated, they were stabilized by self-reassurance. The children of average intelligence had sufficient scope to be excited by the novel stimulation, but not enough to be automatically stabilized by those items which should guarantee confidence in the matter as a whole. The highest degree of disturbability was located between 90 and 100 IQ among the various ranges of IQ studied.

The application of paper and pencil inventories and questionnaires to persons of low degrees of intelligence is not feasible because of their very limited capacity for literacy. Hence there exists little in the literature of personality inventories that relates to very inferior intellect. Blair (8) has recently summarized the literature on *interests,* and has presented data from questionnaires showing that high-school pupils of inferior intelligence differ from those of superior intelligence, in their interests in a variety of fields.

Durost (30) has shown that an interest in collecting may be manifest in some individuals in the IQ range below 90. Simmons (86) found that schoolchildren testing below 90 IQ are more easily influenced by others in making decisions than are children who test above 130 IQ. Hartshorne and May (41) discovered that the attempt to cheat and deceive for the purpose of winning a score is greatest among schoolchildren of subnormal intelligence. Haggerty (39) found that 'undesirable behavior,' as defined by teachers, centers overwhelmingly among pupils who grade under 90 IQ.

The development of delinquent personality is discussed by Williams in Chapter X of this Yearbook. Attention is here called to his summary statement: "The typical delinquent, then, seems to be of the dull-normal classification." (The mean of recent findings of the central tendency of IQ for delinquents falls at 85.)

Although Williams finds the intelligence scores of delinquents (apprehended breakers of the law) center at a low point, his final position is that no *causal* connection has as yet been demonstrated between lack of intelligence and delinquent behavior. "It may well be," he says, "that under-privilege, while adversely affecting intelligence through inadequate stimulation, is at the same time disproportionately productive of conditions basic to delinquency."

This conclusion — that there is as yet no evidence that lack of intelligence in itself predisposes to error of the sort that constitutes infraction of law — is perhaps not altogether ruled out by crucial experiment; but certainly formal logic and ordinary inference unite in throwing their weight against it. Lack of intelligence leads to all *other* sorts of error; why not to errors of the particular sort that break the law? This is the weight of common sense. Again, wherever delinquents are found, a low level of IQ is found to prevail; and at the same time, a low level of environment in the background. Since lack of intelligence is by definition proneness to error, a very reasonable syllogism runs: low intelligence is a source of error; error leads to poor environment; error leads to infraction of law; therefore, low intelligence causes both poor environment and delinquency, through its characteristic proneness to error. This is the weight of formal logic.

This united weight of common sense and logic is implicit in the statement by Doll (24):

> The feebleminded as a class are not aggressively antisocial, nor are they aggressively promiscuous in their sex relations. On the contrary, the feebleminded are passive and timid as a class. Their suggestibility, their ingenuousness and their lack of foresight make them easy victims of social circumstances. They do not seek unsocial lives. Rather, their lives are made unsocial because of their inability to cope with the difficulties of modern social life. (p. 38)

One more point in connection with the presentation by Williams calls for comment here. Williams notes that the percentage of low intelligence ratings found in studies of more than a decade ago has been much reduced in more recent surveys, and he believes that this systematic shift in findings is due largely to using an 'adult age' of 14 or 15, instead of the previously used 16, in calculating IQ. As for this reasoning, it would be correct to say that the earlier percentages were probably the more nearly accurate, insofar as any change in the concept of 'adult age' enters in, because study of what divisor should be used in calculating IQ for adults and adolescents seems to show that 16 years is really 'best' (15, 79).

It should be noted, in passing, as a matter of measurement technique bearing upon problems of representation, that the IQ falls slightly, year by year, when the same subnormal individuals are consecutively measured (13, 61), and that the use of the PC (44, 45) shows greater constancy for subnormals.

Questions of percentage representation of unintelligent persons in various groups will not be finally closed until the methods of measurement become more refined.

To say that a large majority of delinquent personalities are found to belong to groups in which the IQ is inferior is not to say that a large majority, or any majority, of inferior deviates become delinquent. Studies show (26, 29) that the typical person of IQ below 90 is 'at large' and occupied in some entirely legal mode of life.

7. The Concept of Feeblemindedness

Within the past decade, the concept of feeblemindedness has changed but little. Feeblemindedness is, as in the decade of the 1920's, still defined in terms of IQ and of centile status. For practical purposes, persons who test below 70 IQ are called 'feebleminded.' IQ, as understood in this definition, is usually a Stanford-Binet IQ, though not always specified as such.

In New York State in 1930, school regulations specified 75 IQ, instead of 70 IQ, as a minimum for placement in ungraded classes for mental defectives (95).

8. The Predictability of Intellectual Inferiority

The preponderance of evidence is that IQ cannot be predicted under school age by methods at present available (3, 57, 84). There are several reasons for this: first, available scales for preschool children have not been standardized on unselected samples, as it is an insoluble problem how to obtain unselected samples of children who are not under public authority; second, children of preschool age are subject to negativism to a degree that tends to reduce test scores (81, 83); third, intelligence is not now, and may never be, measurable until a stage of development later than infancy has been reached when there are present certain characteristic functions, notably those resident in language. Like the teeth, these elements of the person, though implicit in the organism, do not appear till a later stage of development, and therefore cannot be measured until then.

Though the view that the IQ cannot be determined within a small probable error until at least six years of age is the view held by a majority among those who contribute to the literature, it is not a unanimous view. Wellman (101) and Stoddard (100) are the leaders of a minority of experts who believe that IQ's obtained in infancy, and by methods now available, are meaningful in terms of later years. That this is the view held by these investigators is sufficiently clear from their contributions to this Yearbook.

Wellman interprets changes in IQ's obtained under three or four years of age to represent *real* changes in the mental status of the children retested. Since this interpretation has been sufficiently discussed by collaborators on this Yearbook, it is only necessary to refer to this discussion here. The instability of IQ, found by Wellman and other investigators in Iowa, is no doubt closely connected with the fact that their results rest upon the validity of initial tests made when the children were under forty-eight months of age.

The weight of expert opinion, however, is preponderantly that inferior deviates cannot be reliably classified as such until the age of five or six years — with the exception of those extreme degrees of deviation later to be classified as idiocy or imbecility, and even these extreme degrees may sometimes be mistakenly classified by tests made in the early preschool years.

Errors of classification are likely to be in the direction of underestimation, thus classifying as inferior deviates, children who are really of good intelligence. This tendency to underestimation arises first because available 'baby tests' are standardized on selected children above the average, and second because the characteristic negativism and flightiness of very young children are strong factors working toward a reduction of score, whereas there is no counteracting factor in their case, working toward spuriously high scores (84).

At and after school age, however, mental inferiority becomes highly predictable on the basis of intelligence tests now available. Subsequent performance for groups relates closely to IQ after school age is reached, which means that the prediction is highly reliable for the great majority of individuals. In a few instances there will be a large error, either plus or minus, in the first test, with a large contrary error in the subsequent test, producing a discrepancy of more than ten points. The probable error of a testing after school age is reached is not more than $\pm$ 5 points by Stanford-Binet.

As years go on, the difference between two separate tests on the same individual increases somewhat, but in the great majority of cases inferior deviates maintain their approximate centile status *when measured after school age has been reached,* and this whether environment has been modified or not, as appears in the contributions to Part II of this Yearbook. In a school devoted to inferior deviates, Pritchard and others found no appreciable rise in IQ as a result of two years of special education. This confirms the results of Lamson (63) and others.

9. The Problem of Secondary Cases

In those cases where intelligence originally good has been degraded by accident or disease, especially in cases classified as due to endocrine malfunction, many interesting problems remain to be solved. Inasmuch as physiological conditions in relation to intelligence are discussed in Chapter IV, these matters will be omitted here.

III. Superior Deviates

Theoretically, we should expect general laws to govern individual differences, which would imply that for superior deviates the converse will obtain of all facts found to hold for inferior deviates (mankind being all one species throughout all degrees of intelligence manifested by *homo sapiens*). It is fascinating to see how truly this theoretical expectation is borne out in facts, as they have been obtained slowly, one by one. The picture of highly intelligent deviates as organisms is a recent construct, as compared with the picture of inferior deviates. The latter were studied first, because they gave trouble at school and in society, which superior deviates did not give.

1. The Census of Superior Deviates

Galton was the first to suggest (33) and Terman was the first to prove (91) that the distribution of general ability in a large, unselected sample is symmetrical, and that for every individual testing at a given point below the mean, there is another testing at the analogous point above.

The populations most easily accessible for testing are those of schoolchildren, and available censuses of the gifted are mostly based on schoolchildren.

2. Frequency as Related to Sex

Surveys of superior deviates, by means of intelligence tests, planned for the express purpose of comparing the sexes, have not been carried out. Findings with regard to the frequency of the two sexes have always been incidental to some other main situation. Here, again, the differential action of social forces obtains, and aggregations spontaneously existing practically never contain precisely the same sampling of both sexes. How to obtain exactly the same intellectual sample of both sexes is in itself a difficult statistical problem.

Terman (92), in 1925, found more boys than girls in the samples collected by him on the Pacific Coast. These samples were gathered without any primary intention to compare the sexes. For instance, Terman's sample of younger children includes more than a hundred who were *volunteered* for testing. *Volunteered* children always have among them a large majority of boys, because parents consider it more important to have boys tested. This is a matter of general observation, but precise evidence of it is to be found in the work of Rust (83).

As Terman's plans for identification became more and more systematic and independent of personal nominations, the preponderance of boys steadily declined in his sample of younger children. However, in his sample of adolescents in the high school, the preponderance of boys was very great.

Recently Witty (104) has offered statistics that differ markedly from those found by Terman in the high schools. Witty found .32 percent of boys and .35 percent of girls testing above 140 IQ in the high schools he surveyed. Among Negro pupils in Grades III to VIII, Witty found girls more than twice as numerous as boys in the range above 120 IQ.

The comparative frequency of the sexes in the higher ranges of intellectual ability thus remains a matter for study. The question cannot be closed except by studies planned explicitly to make this comparison, which will guard rigidly against all the strong and very subtle forces making for differential selection.

3. Frequency as Related to Ethnic Stock

From what has already been said about the ethnic stocks in the United States that yield a large percentage of inferior deviates, the inference arises that these same stocks will probably contribute a small percentage, or none at all, to the census of the highly intelligent. This inference is correct. According to mental surveys so far made, the American Indian, the Negro, the Mexican, and the Sicilian yield few superior deviates (34, 35). To this may be added that the Portuguese in California contributed few or no children who tested above 140 IQ to Terman's sample (92).

On the other hand, certain ethnic stocks contribute more than their share of gifted children to the schools of the United States. Such stocks, so far recorded, are the Jews, the Chinese, the Danish, Swedish, and Norwegian, the German (34, 92), and the English and Scottish (103).

A number of ethnic-stock, and also national-origin, groups in the United States have not been surveyed from this point of view. For instance, the children of Irish descent are largely in parochial schools, and have not been available for surveys conducted by public schools. The Welsh in Pennsylvania have not been recorded; and numerous other distinguishable immigrant groups remain to be studied.

Although some national-origin groups contribute very little to the top centile on the distribution of American IQ, none of those so far studied fails wholly to contribute. In other words, superior deviates apparently may arise from any ethnic stock at present existing in the United States — a feature of the situation as important as are the differences in central tendency among stocks.

Witty and Jenkins (105) reported the case history of a thoroughbred Negro girl of 200 IQ, in the city of Chicago. The same investigators identified many other Negro children testing in high centiles of the population (106). In Rapid Learner Classes, at Public School 500, Manhattan, Speyer School, are four Negro children, testing respectively at 130, 140, 155, and 158 IQ. These children maintain themselves intellectually in the work of the school indistinguishably from white children of equal IQ, year after year.

Crump (19), in testing fifty children from each of the five civilized tribes in the Oklahoma reservation schools, found the mean IQ's to rate as follows: Choctaw, 98; Chickasaw, 87; Seminole, 84; Creek, 96; Cherokee, 87. In three of these samples were found children who tested in the top centile for white children. Three Choctaws and one Cherokee tested at or above 140 IQ.

Terman (92) reports the child of a Mexican laborer and a few Negro children who tested above 140 IQ.

The proved occurrence of a few very superior deviates in groups generally low in test scores, and also low in social esteem and opportunity, presents an intricate and interesting problem for educators. How to educate and how to encourage these children is an ethical, as well as an educational, problem.

The only summary remarks that can at present be offered on this topic are, first, that investigation of the past two decades has uncovered great differences in the frequency with which various distinguishable and namable ethnic and national-origin groups in the United States contribute children testing in the top centile of intelligence, and, second, that further research in this field is urgently needed.

What is found concerning superior deviation among samples located in the United States tells nothing, of course, about the quality or status of mother populations from which those samples have been derived. Immigration into the United States has been motivated differently at different times, and for different peoples. The qualities of mother populations must be determined on home ground.

A literature is slowly accumulating, which will in time tell whether more high deviates exist among some peoples than among others. It is interesting to note from data already furnished for large samples of children in England and Scotland that the same statistical findings seem to hold for superior deviation as in the United States (27, 28), in that about 1 percent test at or above 130 IQ. This fact is also brought out by Rusk in Part II of this Yearbook, Chapter XVIII.

Intimately connected with the distribution of superior deviates as found by test is the study of the distribution of notable adults. J. McKeen Cattell (12) and Visher (97, 98) have made further contributions to this inquiry within recent years.

In 1927 Cattell found that scientific leaders in the United States were being born decreasingly in New England, and increasingly in the Middle West. Visher showed the same tendency for notables in ten various categories. Both investigators recorded a paucity of birthplaces of notables for the Southern states.

Visher (98) found that disproportionately large yields of American notables of the 1920's were made by the Quakers, Scotch, Germans of the immigration of 1848, and Yankees.

> But according to much evidence, by far the largest group of American leaders are descendants of emigrants from East Anglia, England, the district which, according to Havelock Ellis ('British Genius'), yielded relatively many eminent British and most of the Puritans of New England, the Quakers, and the Cavaliers of Virginia. (p. 177)

Visher suggests that,

> The westward spread of the descendants of the Puritans out from New England clearly helps explain the geographical contrasts in the yield of notables. Conversely, their partial submergence numerically by less valuable stocks largely produces the decline in the yield of notables in proportion to population in New England and various more western areas. (p. 177)

Aretz, Gutherie, and Waldman (4) found "a steady northern trend of nativity of the notable persons included in the *Encyclopaedia Britannica*." This fact they do not interpret, preferring to leave their study purely descriptive.

4. Frequency as Related to Socio-economic Status

It has been shown that children who yield low intelligence quotients come disproportionately from families of low socio-economic status. The converse holds for children who test high. All surveys have agreed in finding that a large majority of children testing at or above 140 IQ have been fathered by persons in classes I to III on occupational scales devised by economists; and that classes IV, V, and VI yield very few of such children. The finding by Terman (92) that fathers classified as 'professional, semi-professional, and business' yielded 81.4 percent of his large sample of children above 140 IQ is typical; in comparison 'common labor' yielded only 0.13 percent. The professional group of fathers yielded 1003 percent of their quota (number of offspring to be expected if all occupational groups were alike).

The median income reported for Terman's families of gifted children was $3,333 per annum, between two and three times as great as that for the population at large in California, in 1922–1923; and yet not constituting wealth.

These findings are typical, and they are in agreement with the facts about socio-economic origins of American notables. Visher (98) states that American notables of the 1920's were born into "the so-called upper classes," in a very large majority of cases. The several geographic sections of America varied in their yield of notables only roughly in accord with their relative numbers of such classes.

Furthermore, the finding is the same for origin of the most eminent persons of history. Cox (18) found that origin in families of high standing and economic security was typical of the 312 eminent persons whom she selected for study of their childhood.

These findings will doubtless be confirmed wherever gifted children may be studied. However, it is necessary to emphasize the fact that a substantial and important minority of superior deviates originate in poverty, in families where all means are lacking for appropriate development (88). At the Speyer School,[1] a study is at present under way to show the socio-economic condition of pupils in Rapid Learner Classes. Most of the families of these children who test from 130 to 200 IQ are in poverty, in some cases amounting to dependency. The series of misfortunes leading to this condition is very interesting. Death or default on the part of the father is the chief causal factor. Many are children of widows, whose fathers were mental workers and died young, before the accumulation of money was possible. Each story is unique, but the total shows that many of the children who test at or above 140 IQ live in want with no likelihood of family support for higher education. A substantial number of children whose potential achievement is demonstrated by an IQ above 170 have no prospect of family support for their subsequent development. The presence of

[1] In New York City, families in the higher income brackets are apt to patronize private or parochial schools.

these children of great promise who are sorely handicapped by poverty or near-poverty constitutes a responsibility for education and for government (42, 53).

5. Qualities Correlated with Superior Intelligence

Just as inferior intellectual deviates tend as a group to be undersized, lacking in vitality, and exhibiting an unusual number of physical defects, so superior deviates are conversely characterized. All studies of the qualities of highly intelligent individuals show a general quality of the organism (46) that places these persons above the average in stature (49, 92). This accords with the findings that establish for measurements of general populations positive correlations of about .20 between intelligence and the magnitudes of physique (58). The greater size of the highly intelligent as a group is found at all periods of development investigated, and for the same individuals followed consecutively from childhood through adolescence (49). Stature and weight are reliably superior for the highly intelligent, as a group. It remains to investigate the hypothesis that the extremities of the highly intelligent are proportionately smaller and more delicately formed than those of the generality. This hypothesis is offered by the present writer, and applies to hands, feet, ears, and nose.

New investigations of the neuromuscular capacity of superior intellectual deviates have not appeared. The existing studies show them to be superior in performances that do not involve the lifting of body-weight (58, 71), though there is much overlapping with the generality.

By the use of photographs, beauty of the face has been studied by Hollingworth (51), with the result that adolescents of 135 to 200 IQ were consistently rated higher for beauty than adolescents of average intelligence.

Greater longevity for persons above 135 IQ may be indicated by the fact that, of 56 such persons, followed from the age of 8 to the age of 26 by the present writer, in New York City, only one has died (age 24). This compares favorably with the expectation from mortality tables.

New data and observations bearing on the temperament and character of the highly intelligent have appeared. Conklin (17) has shown that pupils of IQ above 130 may fail in their studies in the high school through what are, apparently, as yet unanalyzable traits of personality, a ' quid ' that is constitutional in its nature. Laycock (64) finds that superior deviates make superior adjustments. Regensburg (80) also studied success and failure; and Terman contributes herewith an account of success and failure in the subsequent histories of the children first reported by him in 1925 (92).

Hollingworth and Rust (56), applying the method of Bernreuter, found adolescents who had rated in childhood at 135 to 200 IQ to be less neurotic, more self-sufficient, and more dominant in attitude as a group than those on which the comparable norms were established. The study by Collmann (16), already cited, found children above 125 IQ to be more stable than the generality in respect to stimulation under the psychogalvanometer.

In short, the older studies are confirmed by the new, in reporting that children from 130 to about 150 IQ as a group showed superior adjustment of all sorts on the basis of character and temperament. These children are close enough to the generality to have a 'common sense' with those whom they meet, and are sufficiently frequent in occurrence to find others like or nearly like themselves. Beyond about 160 IQ, however, the picture begins to change. Children begin at about 160 IQ to show characteristic difficulties of adjustment, and in the ranges above 170 IQ the tendency to become isolated, with all that this implies, is strongly in evidence (10, 50, 52). More will be said of this matter in discussing the concept of genius. From observation of deviates, the concept of an *optimal range* of IQ, as distinguished from maximum range, formulates itself (52) — optimal, that is, for social adjustment and success as ordinarily understood.

6. Predictability of Superior Development of Intelligence

Follow-up studies of the past decade show that children selected at and after seven years of age as rating in the top centile of the distribution of Stanford-Binet IQ's maintain their centile status on mental tests with a very high degree of reliability (10, 55, 67). Regression from top centile is rare, and regression to lower than ninety-fifth centile hardly occurs. This is the finding in spite of the fact that end scores in practically all instances have rested on a single test, so that an occasional large error on first test with a large contrary error for the same individual on final test is bound to bring a discrepancy in a predictable number of cases.

The highly intelligent, when followed not with mental tests, but with inquiries into scholastic success and other adjustments, are found to maintain themselves in superior fashion, as a group, though with failure on the part of some individuals (10, 27). This occasional failure indicates the rôle played in adjustment by forces other than intelligence. In other words, intelligence is maintained, as shown by tests, but intelligence alone is not enough for 'success.'

7. The Concept of Genius

Within the decade now ending, a real change has taken place in the concept of genius in relation to IQ. When Terman (91) first undertook to classify individuals under various existing concepts in terms of the Stanford-Binet IQ, he suggested that "genius or near-genius" should be defined as beginning at about 140 IQ. His genetic studies of children at and above 140 IQ he entitled *Genetic Studies of Genius.* The *Dictionary of Psychology,* following this usage, says of the word 'genius': "It has no special technical meaning, but has occasionally been defined as equivalent to an intelligence quotient (IQ) of 140 or above" (99).

It has now become clear from follow-up studies (55, 67) that children who test at 140 IQ are far from genius in any accepted meaning of that term. The median student in many of our highly selected first-rate colleges grades at 150 IQ. The CAVD scores, and Army Alpha scores of the fifty-six children with whom the present writer has been familiar ever since 1922 (67) yielded curves at college age that fit almost exactly over the curves obtained for Barnard College students at those ages. The central tendencies of the two groups coincided exactly, and the mean IQ of our group in childhood was 153. The only difference between the comparative curves was that the college group showed three or four stragglers going below the lower limits of our group, which was cut off sharply at 135 IQ in childhood.

The CAVD scores of the group who had been tested in childhood found Q3 for college graduates in general at about 140 IQ. Probably no one would wish to consider the upper quartile of American college graduates as representing genius. Terman (10) noted the tendency for those from his studied sample who attended Stanford University to fall naturally into the upper half of the student body there.

It is necessary, therefore, to revise the idea that 140 IQ delimits genius. From subsequent observation of the mental products of tested persons, it seems reasonable to suggest that the degree of mental ability involved in producing works of genius falls as far above 140 IQ as the latter falls above the generality. At and above 180 IQ performance begins to appear that corresponds to the lexicographer's idea of genius. Origination, record-breaking, the winning of medals, and like performances are apt to occur in early adulthood, at the level above 180 IQ (67).

The formulation of the concept of genius will ultimately no doubt involve more than the element of IQ. However, IQ will certainly enter as a *sine qua non* of intellectual genius, and it will be a fascinating study to determine what the limits are at which we may begin to expect works of genius and below which no such works are ever found.

In this connection it may be suggested that the IQ's ascribed to historic geniuses in the study by Cox (18) seem much too low to one who has observed tested children develop their subsequent careers over a quarter of a century. It seems likely that judges rating childhood data for these great persons, after twenty years of experience with children of from 140 to 200 IQ, would rate them nearly all above 170 IQ, and a large majority above 180 IQ. It is to be remembered in evaluating these remarks, however, that Cox's raters were instructed to assign the *lowest* IQ that could account for the data of childhood; and that in some cases there were very few data on which to base a rating. In cases of the latter category, it would probably have been better to omit the persons rather than to arrive, for example, at such a conclusion as that the IQ of Faraday in childhood could have been as low as 150, or even as low as 105.

It has been suggested that prediction of works of genius will not be possible from IQ ratings alone. Not all who have the requisite IQ will produce works of genius. Galton's concept of genius (33) as having three essential elements: " The concrete triple event, of ability combined with zeal and with capacity for hard labour " seems more and more valuable. The first element named, ability, is what we may now translate as IQ. That without zeal and energy, ability does not amount to genius seems to be a better concept than to base the concept of genius on IQ alone.

As a matter of fact, probably not more than one in four or five of children testing above 180 IQ enjoys that ' combination of events ' that will lead to eminence. Sex, race, and economic status are important factors. Physical stamina counts heavily. Further, the difficulties of social adjustment are great for such children, sometimes too great to be successfully met (10, 50, 76, 88).

Facts about extremely high deviates are accumulating in the literature from case studies of individuals, as well as from general surveys (8, 35, 68, 105). From these we shall finally learn the scientific facts about the extreme upper limits of intelligence, and the significance of these highest ranges for education and social science.

IV. Miscellaneous Observations

1. Validity of Tests for Deviates

A few contributions have appeared that deal directly with methods of measuring the intelligence of deviates as such. These studies indicate that not all instruments that are valid for distributing the generality are equally valid in the case of deviates. A given test of intelligence may yield the same norm for the population on the whole as another given test, and yet the two may differ greatly as respects validity *for deviates.*

Carroll and Hollingworth (11) showed some years ago that the Herring-Binet is characterized by a systematic error that pulls high deviates toward the center. Children who score above 135 IQ on the Stanford-Binet test score on the average about 17 points lower on the Herring-Binet test. Using performance on Stanford Achievement tests three years after mental testing as a criterion of validity, it clearly appears that Stanford-Binet, and not Herring-Binet, is the valid predictor of performance.

MacMurray (68) compared fifty children who had tested at or above 130 IQ (Stanford-Binet) with fifty children who had tested between 75 to 90 IQ (Stanford-Binet) on the Pintner-Paterson Performance Scale. For these young deviates it was proved that Pintner-Paterson yields spuriously high IQ's for the dull, and spuriously low IQ's for the highly intelligent. Predictions of performance in the literary curriculum of the school will be markedly erroneous for *young deviates* when based on Pintner-Paterson IQ's.

2. The 'Ceiling' in the Case of Superior Deviates

One of the unsolved problems of the testing laboratory continues to be the construction of tests that will prevent the most intelligent individuals from 'going through the ceiling.' This applies especially to adolescents and adults. No tests have as yet been devised that will fully extend and distribute persons of 200 IQ when they reach adolescence and adulthood. CAVD (94) is calculated to 'hold' the highest deviates, but application of CAVD to adults who are known to have tested at or near 200 IQ in childhood shows that these persons elude distribution. Army Alpha, also, does not adequately distribute adults who in childhood reached or exceeded 180 IQ.

Children above 150 IQ begin to pass tests at the highest levels provided on Stanford-Binet, 1916 Revision, when they are about eleven years old, and thus are not thereafter fully extended. At thirteen years of age, such children reach the ceiling of Stanford-Binet, 1937 Revision, as is shown in an unpublished study just completed on high deviates at Speyer School (78). Children above 180 IQ are 'through the ceiling' of Stanford-Binet, 1937 Revision, by the time they are eleven years old.

3. Shifts of IQ

P. Cattell (13), in retesting 1,183 school children of various degrees of intelligence, found a distinct tendency for the low IQ's to become lower, and for the very high IQ's to become higher. "A definite tendency was found for the pupils of high intelligence to gain, and for those of low intelligence to lose as they became older." Lincoln (65, 66), on the other hand, in retesting superior children, found a median drop of four points in IQ. "Over a period of five or more years, there is more loss than gain." The group as a whole remained, however, in its original classification of 'superior.' These discrepancies in results may be incidental to differences in the *ages* at which the two groups of children were tested, and also to differences in amount of superiority. It has just been pointed out that very superior children fall off in IQ by going through the ceiling of Stanford-Binet after eleven years of age.

4. IQ's in Private Schools

Dearborn and Cattell (23) have verified the finding previously announced by other investigators, that where private schools are well established, as on the Atlantic seaboard, they are attended by a very disproportionately large number of children testing above 130 IQ. The median IQ for first-rate private schools seems to be about 120. In areas where such schools exist, there is a strong tendency for the brightest children to be drained off into them from the public schools.

5. Special Talents in the Gifted

As regards special talents of intellectually gifted children, it has been shown by Hollingworth (48) that children who test at or above IQ 135 are distributed as unselected children of their age are distributed in six of Seashore's tests of musical sensitivity. For the six tests as a group the children of high IQ reach a median at the fiftieth percentile. A possible exception is noted in *sense of time,* in which the group of highly intelligent children found their central tendency at the sixtieth percentile.

6. Recognition and Provision for the Gifted in the Schools

A general survey of much importance has been added to the literature by Witty (103), covering high deviates from the Middle West, and supplementing the surveys by Terman (92) on the Pacific Coast and by Hollingworth (55) on the Atlantic Coast. The general facts obtained from all of these surveys are the same: children testing at or above 140 IQ are unrecognized by the school, are unprovided for in mass education, are functioning far below their mental level in the elementary school, yet are maintaining themselves scholastically and socially without giving much trouble to either school or society. They earn high 'marks' without effort and waste one-half or more of their time during the school day either in idleness or in the performance of routine tasks and errands, under the general concept of 'helping the teacher.'

V. Concluding Remarks

It is evident that an exhaustive account of the literature covering intellectual deviates over the past decade of research, minutely documented, would exceed the limits of the present purpose. It has been the intention to refer to studies to which are attached detailed bibliographies; and to summarize briefly knowledge that has accumulated within the past decade, relating such new knowledge to previous findings, insofar as feasible without too much detail. It is believed that from the references cited a competent student of the subject would be able to reconstruct the total literature dealing with intellectual deviates, over the period from 1927, thus becoming informed of all those details that we here necessarily miss.

In closing we would note that there are recently published reviews that may be consulted for original studies of deviation (1).

No attempt has been made to note the literature dealing with educational experimentation with deviates. Scholastic achievement is not treated, except insofar as reference to achievement gives insight into organic quality. The object has been to summarize present knowledge about intellectual deviates as *organisms*, with special reference to research reported since 1927.

With the foregoing review of literature in mind, it is of great interest to consider the summary of recent research on the subsequent history of children testing at or above 140 IQ, presented by Dr. Terman, immediately following, in this chapter.

II

STATUS OF THE CALIFORNIA GIFTED GROUP AT THE END OF SIXTEEN YEARS

LEWIS M. TERMAN
Professor of Psychology
and
MELITA ODEN
Stanford University
Palo Alto, California

I. DESCRIPTION OF THE GROUP

The group referred to is composed of subjects who as children tested at 140 IQ or higher. Allowing liberally for regression of found scores, practically all these children at the time they were tested belonged in the top one percent of the generality of California school children, and a majority of them were in the upper half of the top one percent. The mean IQ of the group was approximately 150.

The subjects were for the most part located in 1922 by a systematic search throughout the larger cities of California. The composition and characteristics of the group have been described in considerable detail in Volume I of *Genetic Studies of Genius*, 1925 (92). Volume III of the same series (10) gives the results of a follow-up study carried out six years after the original group had been located and tested. The most important findings set forth in these two volumes are summarized in the statements that follow:

1. Gifted children come predominantly from family stocks of decidedly superior intellectual endowment and of slightly superior physical endowment.
2. These family stocks have greatly decreased in fecundity during the last two generations and have already reached the point where they are not maintaining themselves.
3. The mean IQ of siblings of children who are in the IQ range above 140 is about 123, or almost exactly what would be expected if the correlation between siblings in the general population were in the neighborhood of .45 or .50.
4. Intellectually gifted children, either because of better endowment or better physical care, or both, are as a group slightly superior to the generality of children in health and physique and tend to remain so.

5. Children above 140 IQ are not as a group characterized by intellectual one-sidedness, emotional instability, lack of sociality or of social adaptability, or other types of maladjusted personality.

6. Indeed, in practically every personality and character trait such children average much better than the general school population.

7. In social-intelligence ratings, social interests, and play activities, gifted children as a group are either normal or superior.

8. In mental masculinity and femininity gifted boys rate on a par with unselected school boys of corresponding age, while gifted girls deviate significantly from the norm of their sex in the direction of greater masculinity.

9. In the character traits measured by the Raubenheimer-Cady tests the typical gifted children of nine years are on a par with unselected children of thirteen or fourteen years.

10. In trait ratings by teachers gifted children show their superiority to the average most of all in intellectual and volitional qualities and least in physical and social traits.

11. In school progress the typical gifted child is accelerated by 14 percent of his age, but in actual mastery of the school subjects (as shown by achievement tests) he is accelerated by more than 40 percent of his age.

12. At the age of ten years there is no correlation between achievement test scores and the number of years gifted children have attended school.

13. As a rule gifted boys maintain or almost maintain their relative superiority to the common run in intelligence, at least through the period of adolescence.

14. Girls somewhat more often than boys show a drop in the IQ as adolescence is approached, or soon thereafter.

15. School achievement as a rule continues through high school and college to be in line with the IQ originally found in 1921–1922.

16. Subject failures in high school are practically never incurred by children of this grade of intelligence.

17. Nearly three-quarters of the total marks earned in high school by gifted girls, and nearly half of those earned by gifted boys, are of A grade.

18. Gifted children of the senior high-school year test on the average above the 90th percentile of the general run of high-school seniors on the Iowa High-School Content Examination, or from 1.5 to 2.0 S.D.'s above the mean of high-school seniors in general.

After 1928 there was no systematic follow-up of the group until 1936. At that time an effort was made to locate as many of the subjects as possible and to secure from them and from their parents or other relatives information regarding their later development. Of the approximately 1,500 subjects on record in our files in 1928, some 1,400, or 93 percent, were located. To each of these was sent a four-page Information Blank (8½ by 10 inches) calling for information on formal schooling since 1928, academic degrees, graduation honors, scholarships and assistantships, college activities, student earnings, plans for further education, occupational history since leaving school, earned income, occupational goals, avocational interests, health, disappointments and failures, profound influences on subject's life, family data, marriage, data regarding the family and education of the spouse, offspring, and certain other topics. A blank of similar size was sent to a parent or other near relative calling for information on the subject's physical and mental health, intellectual promise, special abilities, social traits, desirable and undesirable traits of personality, interest in the opposite sex, occupational success, ambition and drive, marriage, and the subject's parents, sibs, and other near relatives.

By 1938 these blanks had been filled out and returned by more than 96 percent of subjects located; namely, by 755 males and 555 females. The response was similarly good from parents and relatives. This, we believe, sets a new high for response to questionnaire returns. Moreover, the information supplied is believed to be much more accurate than is ordinarily secured by the questionnaire method. This is assured not only by the intelligence and coöperativeness of the subjects and their families, but also by the fact that the subjects knew they were later to be visited and retested by field assistants. In all our contacts with these subjects during more than sixteen years we have encountered not a single instance of gross misstatement of fact. They may not tell all, but what they do tell approximates closely the truth.

The 1922 elementary-school subjects, who composed about three quarters of the entire group, now range throughout the twenties and have a median age of 26. The 1922 high-school subjects range from 28 to 36 and have a median age of 32. The group is therefore sufficiently mature to give some indication as to the kind of men and women gifted children become.

Thus far, 43, or about 3 percent, of the subjects have died: 12 from accidents, 20 from natural causes, 5 by suicide, and 6 from causes unknown. The number who have suffered mental breakdowns serious enough to require hospitalization is less than 1 percent. These figures, with the possible exception of suicides, are perhaps a little lower than for the general population of corresponding age.

II. Their Academic Careers

Nearly 90 percent of the boys and 85 percent of the girls have gone to college. The mean age at entrance was almost exactly 17 years. Of boys entering, 19 out of 20 graduate; of girls, 9 out of 10. The mean score on the Thorndike Intelligence Examination for some 400 of the subjects who entered Stanford University was more than a standard deviation above the mean of the highly selected Stanford population and only a few points below the mean of graduate candidates for the Ph.D. degree in psychology.

Although the gifted are on the average nearly two years younger than their classmates, they are about three times as likely to graduate with honors. Roughly 16 percent of the boys and 20 percent of the girls who graduated from college are elected to Phi Beta Kappa; 12 percent of the boys and 3.5 percent of the girls are elected to Sigma Xi; many others are awarded departmental honors. Two-thirds of the boys and half of the girls who graduate from college take up post-graduate work. Most of the girls continue only to the M.A. degree, but a few have taken the Ph.D. or M.D. degree, and others have taken graduate degrees in law or engineering.

Of nearly 300 boys who had completed their graduate work by 1938, some 12 percent received the Ph.D. degree, 13 percent the M.D., 25 percent a degree in law, 23 percent an M.A. degree, 5 percent a degree in engineering, and 2 percent a degree in architecture.

As undergraduates, 40 percent of the boys and 20 percent of the girls earned half or more of their total expenses. Fourteen percent of the boys were entirely self-supporting. Undergraduate earnings of the boys totaled more than a half million dollars. In the graduate years a third of both sexes were entirely self-supporting. A fourth of the boys were assisted by scholarships, fellowships, or assistantships, their stipends from this source amounting to about $200,000.

In college, as in the high school, the group received more than its proportionate share of class and student-body honors and, except in athletics, engaged more extensively in extra-curricular activities than their older classmates.

Although the large majority made superior records in college, a considerable number did not. About thirty of the subjects were disqualified by low grades, but half of these later completed their college courses. In fact, eight who had been disqualified later successfully completed their work for a graduate degree, three of these for the Ph.D.

In no case does it appear that the poor college record could be attributed to lack of ability. Lack of interest and various kinds of maladjustments seem to have been the most important factors. In several cases low college grades were the result of deliberate neglect of class work in favor of private pursuits or extra-curricular activities. Because prior to college little study had been necessary to secure high grades, many underestimated the amount of work requisite for a satisfactory record in college.

III. Their Choice of Professions

Of the boys who have completed their schooling, half have entered one or another of the professions. Law gets the largest proportion, about 13 percent. Engineering and geology together get 10 percent; medicine, 7 percent; college or university teaching, 8 percent; research in chemistry or physics, 5.5 percent; and religious or social work, 6 percent. Semi-professional and business pursuits account for about a quarter of the total, including 7 percent who are working in motion pictures, radio broadcasting, or music. In 1937, still a depression period, nearly a fifth were working in occupations classifiable as clerical, skilled trades, or retail business. The remainder were widely scattered. There are jazz-band players, Walt Disney artists, movie actors, movie technicians, salesmen, seamen, a rare-stamp dealer, a policeman, and a fox farmer. All seem to have been able to keep off the relief rolls. Less than one percent of those classified as employable were unemployed in 1936.

IV. Their Incomes

The mean earned income of boys is $122 a month at age 21 to 22 and increases in successive two-year periods to $250 at age 30. Highest mean income is earned by Ph.D. and M.B.A. graduates, about $250 a month. Medicine comes next with a mean of $217 a month. Those with M.A. or law degrees average about $200. The mean is about the same ($150) for those who did not go to college and for those who stopped with college graduation. The former, of course, have a temporary advantage in length of experience. Among the highest of the earnings are the following: $12,000 a year by a lawyer still in his twenties, $10,000 by a radio advertising executive aged 27, $7,000 by an investment expert of about the same age, $8,000 by a college professor in his thirties, and $6,000 or more by two musicians.

Of 244 girls who have completed their education and are gainfully employed, about 38 percent are engaged in office work or business; 27 percent are teachers; 6 percent are musicians, dancers, actresses, or entertainers; 5 percent are librarians; 3 percent are artists, decorators, or architects; 2 percent are nurses; and 1.6 percent are physicians. The gainfully employed include a third of the girls who are married. Of those who have completed their education and are unmarried, 12 percent are not gainfully employed. A considerable number of the latter group are engaged in volunteer social work or are making plans for marriage; three are invalids and a few others are awaiting their first job.

The mean earned income of girls working is $120 a month for college graduates, $95 for those who attended college but did not graduate, and $90 for those who did not attend college. The gifted girl may look forward to earning only about half as much in adult life as the gifted boy, perhaps largely because a majority of them are willing to take whatever employment is available to bridge the gap between school and marriage.

V. Their Marriages

Up to two years ago, 46 percent of the boys and 51 percent of the girls had married. Of those married, approximately 8 percent of both sexes had been divorced. Mean age at marriage was a little over twenty-four years for boys and a little less than twenty-three years for girls.

Forty-one percent of the boys and 52 percent of the girls married college graduates. The spouses of 12 percent of the boys and 37 percent of the girls had had one or more years of graduate work. The spouses of 30 percent of the boys and 22 percent of the girls had attended college but had not graduated. The spouses of 6 percent of the boys and 10 percent of the girls had had less than four years of high-school education. Of the college graduates who married into the gifted group, a quarter of the wives and 15 percent of the husbands had received graduation honors.

There have been four intermarriages within the group and a fifth is in prospect. In addition, four of the subjects have married sibs of those included in the group.

Of those married, 40 percent have one or more children and 14 percent have two or more. The total number of offspring is at this date close to 350.

VI. The Effect on Them of Inclusion in the 'Gifted' Group

The subjects were asked what effect their inclusion in the gifted group had had upon them. The answer "no effect" was given by 73 percent of the boys and by 65 percent of the girls. Thirteen percent of the boys and 19 percent of the girls mentioned only favorable effects, 9 percent of the boys and 11 percent of the girls mentioned only unfavorable effects, and the remaining 5 percent of each sex mentioned both favorable and unfavorable effects. The favorable effect most often mentioned was increase in self-confidence; the most unfavorable effects were self-consciousness, conceit, and a tendency to rely upon ability rather than hard work.

VII. Their Moral Record

The moral record of the group has on the whole been well above that of the generality. Three of the boys are known to have had a criminal record. One of the girls spent five years as an inmate of houses of prostitution but has since married a respectable man and settled down. Two or three of the girls are known to have had illegitimate offspring. Several of both sexes have been at one time or another addicted to the excessive use of alcohol.

VIII. Their General Accomplishment

Although a considerable proportion of the subjects have not lived up to their ability, the accomplishments of the group as a whole have been as good as could reasonably have been expected, considering that most of its members are still under thirty years of age. At least half the boys are launched upon promising careers and several are already nationally or internationally known. The group includes novelists, short-story writers, poets, a sculptor, a gifted musical composer, and several scientists who have published important researches. With few exceptions creative intellectual productivity is confined to the males. The females, once they have completed their formal education, seek their life satisfactions in other things than intellectual pursuits.

In appraisal of the group's accomplishments it is necessary to bear in mind the financial and industrial depression of the last ten years. When this began, most of the group were in the high school or college. Several were compelled to forego a college education in order to take over the support of parents or other relatives who had been deprived of their usual sources of income. In other cases, graduate professional training has been made difficult or impossible, the level of occupational placement has been reduced, and the salaries have been adversely affected.

IX. Further Plans for Retesting

A recent grant from the Carnegie Corporation, of New York City, will make possible a more thoroughgoing follow-up of the group during 1939–1941. For this three full-time field assistants have been engaged who will spend a year retesting and personally interviewing the subjects, their parents, and their spouses. Offspring above the age of two years will also be tested. An additional grant from the National Research Council provides for a special study of marital adjustments in the group.

III

CORRELATES OF ADULT ACHIEVEMENT IN THE CALIFORNIA GIFTED GROUP

Lewis M. Terman
Professor of Psychology
and
Melita Oden
Stanford University
Palo Alto, California

I. The Problem

In the preceding portion of this chapter we have summarized briefly the results of a recent follow-up of the California gifted group first studied in 1922. We have there noted a wide range of adult achievement in the group, notwithstanding the fact that as children all tested within the top one percent of the generality in intelligence as measured by the Stanford-Binet scale. The question arises, what facts of childhood or of later development are correlated with greater or less achievement in adult years?

II. The Method Followed

1. Composition of the Two Groups under Comparison

We have sought the answer to this question by an examination of our data for gifted boys only, since a majority of the girls have not gone out for careers, and grades of achievement are for them less clearly defined. In the case of the boys we have compared, on the basis of all the information in our files, the 167 most successful with the 146 least successful. These groups may be thought of roughly as the upper and the lower quarters in terms of success and will be designated as Group A and Group C, respectively.

2. The Criterion of Success

The criterion of success was the extent to which a subject had made use of his superior intellectual ability. Although the criterion is admittedly subjective, the three judges who made the ratings agreed upon certain general principles. In the case of subjects who had completed a graduate university course and had entered one of the professions, academic marks counted heavily as well as professional recognition. Earned income was given little weight, except where it seemed clearly indicative of success as above defined. For example, young lawyers or college instructors with a brilliant academic record but small salary were not penalized for low income. Earnings could not be ignored in the case of those who had gone into business or certain of the semiprofessional pursuits, but the judges were instructed to give weight to this factor only in so far as it appeared that success was attributable to superior intelligence.

With these instructions in mind the judges examined the follow-up records of about 600 of the boys who were 23 years of age or older and classified them into three groups on the basis of rated success: Group A, the upper quarter; Group B, the middle 50 percent; and Group C, the lower quarter. The three judges made their ratings independently, but after this was done they discussed each case on which there had been disagreement and disposed of it by majority vote. The judges included, besides the present authors, Barbara Mayer, research assistant in psychology, Stanford University. The slightly larger number of subjects in Group A than in Group C is accounted for by the greater relative incidence of outstanding success than of decidedly questionable success.

The reader will understand, of course, that our classification of the subjects is entirely tentative. All we can say for it is that on the basis of the evidence available at the time the ratings were made, all the judges agreed that, on any reasonable criterion of success, not only was there no overlapping between Group A and Group C but also that the two were separated by a considerable gap. The future will doubtless bring many unexpected shifts. Ill health, mental breakdown, unhappy marriage, or other misfortune will lower some of the A's to the B or even to the C classification. More often, perhaps, those now in the C group will 'find' themselves and move upward.

The two groups are closely matched for age: the A's range from 23 to 35, with a mean of 27.6; the C's from 23 to 35, with a mean of 27.4. Individuals above 35 and below 23 years of age were arbitrarily eliminated in the interests of homogeneity.

3. Six Illustrative Cases

The following brief descriptions of three cases from each of the two groups illustrate the contrast in achievement to date.

A42 graduated from the high school at 15 years, 10 months. After staying out six months to work, he entered college and completed his work for the Ph.D. degree in biochemistry at 22. He earned three-fourths of his undergraduate expenses and all his graduate expenses, with the aid of fellowships and assistantships. At 23 he became assistant professor and head of the Department of Pharmacology in the medical school of an eastern university. Now at twenty-six, he is an associate professor. He has published numerous scientific articles and won considerable recognition for his research.

A61 exemplifies the conquest of almost insuperable obstacles. He comes from a home of poverty and disorder where the foreign-born and poorly educated parents quarreled constantly until they were finally divorced. Though ability of a high order is found in the family stock, there is also a history of epilepsy, insanity, and feeblemindedness. The subject is lame as a result of infantile paralysis in infancy. Nevertheless, he worked at selling papers, caddying (in spite of his lameness), and as delivery boy, thus contributing throughout his childhood and adolescent years to the family support as well as to his own education. He graduated from the high school at 16, having completed the course in three and one-half years. He graduated from the university at 20, *cum laude,* and was elected to Phi Beta Kappa. At 23 he got his LL.B, second in his class, and was elected to the Order of the Coif. He was entirely self-supporting in college, with the aid of two scholarships. In spite of his outside work he was very prominent in debating and chess both in high school and in college. He has had a number of law articles published and is at present associated with an established law firm.

Although A14 has not yet been fully tried, it was felt that his academic record and great promise entitled him to inclusion in the A group. He finished the high school at 15 with an A record. He graduated from the university with highest honors in chemistry and was elected to both Phi Beta Kappa and Sigma Xi. He was selected by a large utility corporation from a country-wide test of college graduates for a position in chemical research.

The following year he was offered an assistantship in three universities. He selected one of the most outstanding of these and made the same distinguished scholastic record while working for the Ph.D. He has published several scientific articles. He is now, at 25, studying abroad on a travelling fellowship.

C54 completed the eighth grade at 12 and the high school 6 years later after many difficulties and frequent failures. He was expelled from one high school and withdrew from another because of too many demerits. He was a persistent truant and unmanageable at home. After leaving the high school he attended a trade school and became a ship radio operator.

C49 graduated from the high school with honors and as class valedictorian at 15 years, 9 months. He finished college at 19 with a C average, earning about one-fourth of his expenses. After remaining out of college two years because of financial needs and "uncertainty of self" be reëntered for the LL.B, but after getting his law degree, he continued for another semester in order to get a high-school teaching credential. He practiced law for a while without any success, then taught in a high school for two years. He gave that up to take a clerical job for a time and is now a street-car motorman. He is still, at 34, undecided as to his life work.

C81 has the highest IQ and the best scholastic record in the C group, but his lack of drive and poorly adjusted personality have combined to prevent any notable achievement. He graduated from the high school with an A record and as valedictorian of his class, despite two illnesses that kept him out of school for long periods. After high school he worked at clerical jobs for two years. In 1928 he entered the junior college and made an A average in his two years there. He then entered the university and graduated in philosophy with a B average. Unable to find employment, he returned to the university after a year and a half and secured his general secondary teaching credential. Unable to secure a teaching position, he accepted a job as statistical clerk, which he still holds, having risen from $75.00 to $125.00 per month.

We have compared the A and C groups on some two hundred items of information supplied by the early test scores, heredity data, health records, case histories, and trait ratings in the hope that by reading the records backward, so to speak, some light would be thrown on the non-intellectual factors that affect success in school or life. Only the more significant of these comparisons can be presented at this time.

III. The Results of the Comparison

1. Educational and Occupational Status

Let us consider first the educational and occupational status of the two groups. Of the A's, 98 percent entered college and 90 percent graduated. Of the C's, 70 percent entered college and only 50 percent graduated. Of the A's, 72.5 percent have had one or more years of graduate work; of the C's, 19.4 percent.

Among those graduating, 32.5 percent of A's and 1.3 percent of C's were elected to Phi Beta Kappa; 21.8 percent of A's and 2.8 percent of C's to Sigma Xi; 9.9 percent of A's and no C's to both of these honor fraternities. The C's, however, were somewhat handicapped by having to earn more of their college expenses. Earning half or more of their undergraduate expenses were 36.7 percent of the A's and 45.8 percent of the C's. Fifty percent of the A's, but only 10 percent of the C's, received appointments to scholarships, fellowships, or assistantships, and the total stipends from such appointments totaled $116,000 for the A group as compared with $5,600 for the C group.

At the time reports were received in 1936 and 1937, the mean earned income was $288 a month for the A's and $123 for the C's. At ages 23 to 24 the respective means were $210 and $96. The earnings of the A's increased steadily to an average of $336 a month at age 28 and to $500 at age 35, while the mean for the C's did not exceed $162 at any age. It will be recalled that salary was given little weight in the classification of subjects into the A, B, and C groups.

The occupational classification of the A and C groups in 1936–1937 was as follows:

	Percent	
Occupational Class	*A Group*	*C Group*
I. Professional pursuits	70.0	17.1
II. Semiprofessional and higher business	25.6	11.1
III. Clerical, retail business and skilled trades	3.8	33.6
IV. Agriculture	0.6	2.9
V. Minor clerical or semiskilled trades	0.0	25.7
VI. Slightly skilled	0.0	5.7
Unemployed	0.0	3.6

From the foregoing facts it is evident that, at the end of fifteen years from the time the members of these groups were first located, Group A and Group C differ greatly in their educational and occupational achievements. We shall next examine some of the data in our records to see what factors can be discovered that correlate with such differences.

2. Intelligence Scores

With respect to intelligence scores the A group have a slight advantage, but the difference is not reliable. In 1922 the mean Stanford-

Binet IQ was 153.6 for the A's and 150.5 for the C's. For those given the Terman Group Test in 1922 the mean IQ's were respectively 148.5 and 147.5. Six years later the mean Terman Group Test IQ's were 141.2 and 139.1, the drop being attributable partly to statistical regression and partly to lack of ' top ' in the Terman Group Test at this level. The 57 A's and 25 C's who later entered Stanford University earned mean point scores on the Thorndike test of 91.4 and 90.9. Although the differences between means were small, there were eight in the A group and only two in the C group with IQ's of 180 or higher.

3. Achievement Scores and School Marks

The 1922 differences in educational achievement, as measured by the Stanford Achievement Tests, were correspondingly small; the median achievement quotients were 142.5 and 138.8, and the general information quotients 156.8 and 154.4. The A's excelled in play information. The proportion who learned to read before the age of five was the same in each group; namely, 45 percent.

We are forced to conclude that the greater success of the A group cannot to any great extent be accounted for in terms of superior intelligence as measured by the tests used. Both in childhood and in late adolescence the intellectual superiority of the A's is unreliably small. This is also true of school accomplishment at the elementary-school level; the main difference here lies in the fact that on the average the A subjects complete the eighth grade three months younger than the C's.

It is not until the subjects reach the high school that the two groups begin to draw notably apart. Of all high-school marks, 62.5 percent of those earned by A's, but only 28 percent of those earned by C's, are of ' A ' grade. The difference in high-school achievement, as measured by objective tests, lies in the same direction but is smaller than that for class marks. In 1928 the Iowa High School Content Examination was given to thirty-one of the A and to thirty-four of the C group who were then high-school seniors. Rating at the 99th percentile on the Iowa norms for high-school seniors were 45.2 percent of A's and 26.5 percent of C's; at or below the 95th percentile, 3.2 percent of A's and 20.7 percent of C's. These figures indicate that the actual accomplishment of the C's in the high school was probably better than their class marks considered alone would indicate and suggest the possibility that the members of the C group may have been penalized in their class marks because of inferior coöperation or of some sort of personality defects. College records of a number of the C group point in the same direction.

4. Extra-curricular Activities

The difference in achievement at the high-school level is not due to extra-curricular activities. Such activities were in fact nearly twice as prevalent in the A group as in the C group. This would suggest that the A's are socially better adjusted and better adapted for leadership than the C's.

5. Books Read

In view of the superior school record of the A group, it was surprising to find that a record of books read over a period of two months during 1922 showed a reliably larger amount of reading by the C's. The proportion who read more than ten books in the two months was 56 percent for the A's and 69.1 percent for the C's. This difference also may reflect differences in social adjustment or at least in the intensity of social interests.

6. Family Background

No other factor so strikingly differentiates the two groups as the contrast in family background. In 1922, 36.7 percent of the A fathers were in the professional class, as compared with 21.1 percent of the C fathers. Below Class II (semiprofessional or higher business pursuits) were 39.4 percent of A fathers and 58.7 percent of C fathers. The following figures show a marked superiority of A parents in educational background.

	Percent			
Educational Background	*A Group*		*C Group*	
	Fathers	*Mothers*	*Fathers*	*Mothers*
Graduated from college	41.0	18.5	22.6	12.0
Did graduate work	25.3	5.3	14.0	3.8
Eighth grade only	11.6	5.3	19.5	16.3

According to the 1922 reports of parents, a larger proportion of the A group had the advantage of home instruction. Only 29.6 percent of the A's received no home instruction, as compared with 48 percent of the C's. There is no evidence that such instruction materially affected scores on the educational tests, but it may have affected school marks and stimulated educational ambition.

The groups do not differ in number of sibs, but sibs of the A's average 5 points higher in IQ than sibs of the C's. Twice as large a proportion of A sibs as C sibs graduate from college.

7. Racial Differences

There is one marked racial difference in the composition of the two groups: the A group includes nearly three times as large a proportion of Jewish subjects. The figures are 14.3 percent for the A's and 5.3 percent for the C's, as compared with 10.5 percent for the entire California gifted group. The Jewish child is under heavy pressure to succeed, with the result that he achieves more per unit of intelligence than those of other racial stocks. In this connection it may be of interest to note that the A group contains one Negro and one Japanese.

8. Marriages, Divorces, and Offspring

The A's tend to marry somewhat earlier than the C's. Although the distribution of ages is almost exactly the same for the two groups, by 1936–1937, 64.5 percent of the A's and only 48.6 percent of the C's had married. The difference is no doubt related to the greater earning capacity of the A's. In line with this is the fact that 29.5 percent of the C wives, but only 10.7 percent of the A wives, were gainfully employed.

Of those married, almost exactly the same proportion of the two groups had produced offspring; namely, 46 percent. Fifteen percent of the married in both groups reported two children or more. In the long run the A's should surpass the C's in fertility because of their greater tendency to marry young. Favorable also from the eugenic point of view is the fact that the A's appear to marry better: 47.6 percent of their wives are college graduates, as compared with 26.6 percent of the C wives. Nearly 8 percent of C wives, and less than 4 percent of A wives, have not graduated from a high school.

The incidence of divorce or separation is three times as high in the C group as in the A group, the respective proportions being 14.1 percent and 4.6 percent. A difference in the same direction is found for parents of the two groups: of A parents, 11.9 percent are separated or divorced; of C parents, 19.8 percent. It is impossible at present to assess the relative influence of environmental and genetic factors underlying this tendency for marital maladjustment to run in families.

Mental abnormality among near relatives is reported by 41 percent of the C parents and by 33.5 percent of the A parents, but this small difference may be due to greater frankness in the reports by A parents.

9. Quality of Home Environment

At the time of the 1928 follow-up, field assistants visited the homes and rated the quality of home environment for each subject. In the case of 90.3 percent of the A's the home environment was rated as superior. The corresponding figure for the C's was 79 percent.

10. Social Adjustment and Traits of Personality

In 1922 both teachers and parents were asked for information regarding the social adjustments of each gifted child. One or more difficulties of adjustment were mentioned for 23.5 percent of A's and for 33.1 percent of C's. Seven percent of A's and 12 percent of C's were said to be 'teased' by other children or to be regarded by them as 'queer' or 'different.' In 1928 the field assistants rated social adjustment as normal or superior for 87.1 of the A's and for 74.6 percent of the C's. Nervous symptoms of some degree were mentioned by parent or teacher in the case of 35 percent of A's and 43 percent of C's.

In 1922 each child was rated by teacher and parent on twenty-five traits. We have classified the traits into six groups and have computed for each group of traits the mean of the composite parent-teacher ratings. These means were as follows (the smaller means representing higher ratings):

	Mean of Ratings	
Trait Groups	*A Group*	*C Group*
Volitional	4.4	4.9
Moral	4.4	4.7
Emotional	4.3	4.7
Social	5.7	6.0
Intellectual	3.5	3.8
Physical	5.3	5.3
All traits	4.77	5.16

It will be noted that, except for physical traits, the mean ratings are consistently higher for the A group. Note that fifteen years prior to the classification of these subjects on the basis of adult achievement, teachers and parents had been able to discern personality differences that later were found to characterize the two groups. Similar differences were found in the ratings on twelve traits by parents and teachers in 1928.

11. Occupational Preferences

In 1928 a reliably larger proportion of A's than C's expressed a definite occupational preference: 82.8 percent as against 68.5 percent. This might be interpreted to indicate superiority of the A's in self-knowledge, ambition, or integration of personality. This is in contrast to the amount of occupational experience the groups had up to the age of sixteen years: at that time 75 percent of the C's and only 63 percent of the A's had held paid jobs.

12. Masculinity-Femininity Scores

Twenty-two of the A's and 31 of the C's happened to be included in a group who were given the Terman-Miles test of mental masculinity-femininity in 1928. The mean for the A's was +73; for the C's, +64. The difference is about .3 of one sigma. Although not entirely reliable, the higher masculinity scores of the A's are consistent with the total personality pictures suggested by the data summarized in the preceding paragraphs, especially with the greater participation of the A's in extra-curricular activities. On a battery of character and trustworthiness tests given in 1922, the groups did not differ.

13. Pubertal Changes

If we can accept reports of parents as to age when voice changed, puberty was attained .8 of a year earlier by A's than by C's (means 14.3 years and 15.2 years). Voice change as late as sixteen years was reported by 13 percent of the A's and 23.7 percent of the C's. There was no appreciable difference in the mean age reported for learning to walk or to talk.

IV. General Conclusions

Our conclusion is that *for subjects brought up under present-day educational regimes, excess in IQ above 140 or 150 adds little to one's achievement in the early adult years.* This does not mean that potentiality for achievement is the same for IQ's of 140 and 180. We do not believe that it is. The more probable interpretation is that we have not learned how to bring the highest intellectual gifts to normal fruition or how to steer them clear of the dangers that threaten personality development in extreme superdeviates.

The data reviewed indicate that, *above the IQ level of 140, adult success is largely determined by such factors as social adjustment, emotional stability, and drive to accomplish.* To what extent these qualities are the product of environmental influences and to what extent they have a genetic basis, no one can say. There is certainly nothing in our data to warrant giving all or most of the credit for the superior achievement of the A subjects to their vastly superior family environment. The facts themselves are as amenable to interpretation in terms of heredity as in terms of environment. The personality defects so prevalent in the C group may be, and to some extent probably are, genetically related to the educationally and occupationally inferior status of the parents of this group.

References

(1) American Educational Research Association. "Psychological tests and their uses." *Rev. Educ. Research,* 8: 1938, 318–362.

(2) Anastasi, A. *Differential Psychology.* (Macmillan: New York, 1937) 615 pp.

(3) Anderson, J. E. "The prediction of terminal intelligence from infant and preschool tests." (Unpublished study.)

(4) Aretz, C. W., Gutherie, M. A., and Waldman, J. L. *The Northward Trend of Nativity of Notable Leaders of Civilization.* (Temple University: Philadelphia, 1931) 46 pp.

(5) Baller, W. R. "A study of the present social status of a group of adults who, when they were in elementary schools, were classified as mentally deficient." *Genet. Psychol. Monog.,* 18: 1936, 165–244.

(6) Bennett, A. *A Comparative Study of Subnormal Children in Elementary Schools.* (Teachers College, Columbia University, Contributions to Education, No. 510: New York, 1932) 81 pp.

(7) Bere, M. *A Comparative Study of the Mental Capacity of Children of Foreign Parentage.* (Teachers College, Columbia University, Contributions to Education, No. 154; New York, 1924) 105 pp.

(8) Blair, G. M. *Mentally Superior and Inferior Children in the Junior and Senior High School; A Comparative Study of their Backgrounds, Interests, and Ambitions.* (Teachers College, Columbia University, Contributions to Education, No. 766: New York, 1938) 87 pp.

(9) Brander, T. "Existerer det ett samband mellan födelsevikt och intelligenskvot hos för tidigt födda barn?" *Finska Läk-Sällsk. Handl.,* 77: 1935, 69–80.

(10) Burks, B. S., Jensen, D. W., and Terman, L. M. *Genetic Studies of Genius. III. The Promise of Youth.* (Stanford University Press: Stanford, California, 1930) 508 pp.

(11) Carroll, H. A., and Hollingworth, Leta S. "The systematic error of

Herring-Binet in rating gifted children." *Jour. Educ. Psychol.*, 21: 1930, 1–11.

(12) CATTELL, J. McK. "The origin and distribution of scientific men." *Science*, 66: 1927, 513–516.

(13) CATTELL, P. "Constant changes in the Stanford-Binet IQ." *Jour. Educ. Psychol.*, 22: 1931, 544–550.

(14) CATTELL, R. B. *Crooked Personalities in Childhood and After.* (D. Appleton-Century: New York, 1938) 215 pp.

(15) CHIPMAN, C. E. "The constancy of the intelligence quotient in mental defectives." *Psychological Clinic*, 18: 1929–1930, 103–111.

(16) COLLMAN, R. D. *The Psychogalvanic Reactions of Normal and Exceptional Children.* (Teachers College, Columbia University, Contributions to Education, No. 469; New York, 1931) 88 pp.

(17) CONKLIN, A. M. *Failures of Highly Intelligent Pupils, A Study of Behavior by Means of a Control Group.* (Dissertation, Teachers College, Columbia University, New York, 1939.) To be published in Teachers College, Columbia University, Contributions to Education.

(18) COX, C. M. *Genetic Studies of Genius. II. The Early Mental Traits of Three Hundred Geniuses.* (Stanford University Press: Stanford, California, 1926) 842 pp.

(19) CRUMP, B. L. *Educability of Indian Children in Reservation Schools.* (Southeastern State Teachers College, Contributions to Education, No. 3: Durant, Okla., 1932) 58 pp.

(20) DAVENPORT, C. B., and MINOGUE, B. M. "The intelligence quotient and the physical quotient: Their fluctuation and intercorrelation." *Human Biology*, 2: 1930, 473–507.

(21) DAYTON, N. A. "Correlation between intelligence and physical condition in 14,176 retarded school children." *Med. Jour. and Rec.*, 132: 1930, 222–224.

(22) DAYTON, N. A., and others. "Mortality and expectation of life in mental deficiency in Massachusetts: Analysis of 14-year period 1917–1930." *New Eng. Jour. Med.*, 206: 1932, 555–570; 616–631.

(23) DEARBORN, W. F., and CATTELL, P. "The intelligence and achievement of private-school pupils." *Jour. Educ. Psychol.*, 21: 1930, 197–211.

(24) DOLL, E. A. "Social adjustment of the mentally subnormal." *Jour. Educ. Res.*, 28: 1934, 36–43.

(25) DOLL, E. A. "The Vineland Social Maturity Scale." Revised Edition. *Training School Bull.*, 32: 1935, 1–7.

(26) DOLL, E. A., and MCKAY, B. E. "The social competence of special-class children." *Jour. Educ. Res.*, 31: 1937, 90–106.

(27) DUFF, J. F. "Children of high intelligence. A follow-up inquiry." *Brit. Jour. Psychol.*, 29: 1939, 413–439.

(28) DUFF, J. F., and THOMSON, G. "The social and geographical distribution of intelligence in Northumberland." *Brit. Jour. Psychol.*, 14: 1923, 192–198.

(29) DUNLOP, F. S. *Subsequent Careers of Non-Academic Boys.* (Privately printed: Ottawa, Canada, 1935) 93 pp.

(30) DUROST, W. N. *Children's Collecting Activity Related to Social Factors.* (Teachers College, Columbia University, Contributions to Education, No. 535: New York, 1932) 115 pp.

(31) ELWOOD, M. "A descriptive study of a gifted child." *Pittsburgh Schools*, 12: 1938, 169–173.

(32) FRAZER, J. A. (See Reference 81a.)

(33) GALTON, F. *Hereditary Genius.* (Macmillan: London, 1869 and 1914) 379 pp.

(34) GARTH, T. H. *Race Psychology.* (Whittlesey House, McGraw-Hill: New York, 1931) 260 pp.

(35) GOLDBERG, S. "A clinical study of K., IQ 196." *Jour. Applied Psychol.*, 18: 1934, 550–560.

(36) GOODENOUGH, F. L. "Racial differences in the intelligence of school children." *Jour. Exper. Psychol.*, 9: 1926, 388–397.

(37) GOODENOUGH, F. L. "The relation of the intelligence of preschool children to the occupation of their fathers." *Amer. Jour. Psychol.*, 40: 1928, 284–302.

(38) GORDON, H. *Mental and Scholastic Tests among Retarded Children.* (Board of Education, Education Pamphlet No. 44: London, 1923) 92 pp.

(39) HAGGERTY, M. E. "The incidence of undesirable behavior in public-school children." *Jour. Educ. Res.*, 12: 1925, 102–122.

(40) HAGGERTY, M. E., and NASH, H. B. "Mental capacity of children and paternal occupation." *Jour. Educ. Psychol.*, 15: 1924, 559–572.

(41) HARTSHORNE, H., and MAY, M. *Studies in Deceit.* (Macmillan: New York, 1928. 2 vols.) 414, 306 pp.

(42) HARVARD COLLEGE. "The development of the Harvard national scholarship plan." (Report of the Dean of Harvard College to the President of the University for the Academic Year, 1936–1937)

(43) HEALY, W., and BRONNER, A. F. *New Light on Delinquency and its Treatment.* (Yale University Press: New Haven, Conn., 1936) 226 pp.

(44) HEINIS, H. "A personal constant." *Jour. Educ. Psychol.*, 17: 1926, 163–186.

(45) HILDEN, A. H. "A comparative study of the intelligence quotient and Heinis' personal constant." *Jour. Applied Psychol.*, 17: 1933, 355–375.

(46) HOLLINGWORTH, H. L. *Mental Growth and Decline.* (D. Appleton-Century: New York, 1927) 396 pp.

(47) HOLLINGWORTH, L. S. "Differential action upon the sexes of forces that tend to segregate the feebleminded." *Jour. Abnorm. and Soc. Psychol.*, 17: 1922–1923, 35–57.

(48) HOLLINGWORTH, L. S. "Musical sensitivity of children who test above 135 IQ. (Stanford-Binet)." *Jour. Educ. Psychol.*, 17: 1926, 95–109.

(49) HOLLINGWORTH, L. S. "Do intellectually gifted children grow toward mediocrity in stature?" *Ped. Sem., and Jour. Genet. Psychol.*, 37: 1930, 345–360.

(50) HOLLINGWORTH, L. S. "The child of very superior intelligence as a special problem in social adjustment." *Mental Hygiene*, 15: 1931, 3–16.

(51) HOLLINGWORTH, L. S. "The comparative beauty of the faces of highly intelligent adolescents." *Ped. Sem., and Jour. Genet. Psychol.*, 47: 1935, 268–281.

(52) HOLLINGWORTH, L. S. "Development of personality in highly intelligent children." *Yearbook Nat. Element. School Prin.*, 15: 1936, 272–281.

(53) HOLLINGWORTH, L. S. "Some suggestions on scholarships." *Ind. Jour. Columbia Univ.*, 4: 1936, 4.

(54) Hollingworth, L. S., and COBB, M. V. "Children clustering at 165 IQ and

children clustering at 146 IQ compared for three years in achievement." *Twenty-Seventh Yearbook, Part II,* of this Society, 1928, 3–33.

(55) HOLLINGWORTH, L. S., and KAUNITZ, R. M. "The centile status of gifted children at maturity." *Ped. Sem., and Jour. Genet. Psychol.,* 45: 1934, 106–120.

(56) HOLLINGWORTH, L. S., and RUST, M. M. "Application of the Bernreuter inventory of personality to highly intelligent adolescents." *Jour. Psychol.,* 4: 1937, 287–293.

(57) HONZIK, M. P. "The constancy of mental test performance during the preschool period." *Ped. Sem., and Jour. Genet. Psychol.,* 52: 1938, 285–302.

(58) HONZIK, M. P., and JONES, H. E. "Mental-physical relationships during the preschool period." *Jour. Exper. Educ.,* 6: 1937–1938, 139–146.

(59) KEYS, N., and NATHAN, J. M. "Occupations for the mentally handicapped." *Jour. Applied Psychol.,* 16: 1932, 497–511.

(60) KLINEBERG, O. *Negro Intelligence and Selective Migration.* (Columbia University Press: New York, 1935) 66 pp.

(61) KUHLMANN, F. "Results of repeated mental reëxaminations of 639 feeble-minded over a period of ten years." *Jour. Applied Psychol.,* 5: 1921, 195–224.

(62) LAMSON, E. E. *A Study of Young Gifted Children in Senior High School.* (Teachers College, Columbia University, Contributions to Education, No. 424: New York, 1930) 117 pp.

(63) LAMSON, E. E. "To what extent are intelligence quotients increased by children who participate in a rich vital school curriculum?" *Jour. Educ. Psychol.,* 29: 1938, 67–70.

(64) LAYCOCK, S. R. "Adjustments of superior and inferior school children." *Jour. Soc. Psychol.,* 4: 1933, 353–366.

(65) LINCOLN, E. A. "The Stanford-Binet IQ changes of superior children." *School and Soc.,* 41: 1935, 519–520.

(66) LINCOLN, E. A. "A study of changes in the intelligence quotients of superior children." *Jour. Educ. Research,* 29: 1935, 272–275.

(67) LORGE, I., and HOLLINGWORTH, L. S. "Adult status of highly intelligent children." *Ped. Sem., and Jour. Genet. Psychol.,* 49: 1936, 215–226.

(68) MACMURRAY, D. "A comparison of the intelligence of gifted children and of dull-normal children measured by the Pintner-Paterson Scale, as against the Stanford-Binet Scale." *Jour. Psychol.,* 4: 1937, 273–280.

(69) MCELWEE, E. W. "Seymour, a boy with 192 IQ." *Jour. Juv. Research,* 18: 1934, 28–35.

(70) MCNEMAR, Q., and TERMAN, L. M. "Sex differences in variational tendency." *Genet. Psychol. Monog.,* 18: 1936, 1–66.

(71) MONAHAN, J. E., and HOLLINGWORTH, L. S. "The neuro-muscular capacity of children who test above 135 IQ." *Jour. Educ. Psychol.,* 18: 1927, 88–96.

(72) MORRISON, B. M. *A Study of the Major Emotions in Persons of Defective Intelligence.* (University of California Press: Berkeley, 1924. University of California publications in psychology, Vol. 3, No. 3) 145 pp.

(73) NATIONAL COMMITTEE ON COÖRDINATION IN SECONDARY EDUCATION. "Education of Gifted Children in Secondary Schools. Report adopted at Cleve-

land, Feb. 27, 1939." (Accepted for publication in the *Jour. of Educ. Sociology*, Vol. 12.)

(74 NATIONAL EDUCATION ASSOCIATION OF THE U. S. A., COMMITTEE ON SOCIAL-ECONOMIC GOALS OF AMERICA. *Implications of Social-Economic Goals for Education.* (Washington, D. C., 1937) 126 pp.

(75) NEFF, W. S. "Socioeconomic status and intelligence: A critical survey." *Psycho. Bull.*, 35: 1938, 727–757.

(76) NEVILL, E. M. "Brilliant children: with special reference to their particular difficulties." *Brit. Jour. Educ. Psychol.*, 7: 1937, 247–258.

(77) PATERSON, D. G. *Physique and Intellect.* (D. Appleton-Century: New York, 1930) 304 pp.

(78) PRITCHARD, M. C., and others. " A comparison of the two Stanford Revisions of the Binet-Simon Scale as respects validity for high deviates." (Unpublished study.)

(79) RAPPAPORT, M. E. "The selection of the intelligence quotient divisor for clinical cases between fourteen and nineteen years of age." *Jour. Educ. Psychol.*, 25: 1934, 101–114.

(80) REGENSBURG, J. " Studies of educational success and failure in supernormal children." *Arch. Psychol.*, 20: 1931, 1–150.

(81) REYNOLDS, M. M. *Negativism of Preschool Children.* (Teachers College, Columbia University, Contributions to Education, No. 288, New York, 1928) 126 pp.

(82) RUSSELL, B. *Icarus: Or The Future of Science.* (Dutton: New York, 1926) 64 pp.

(83) RUST, M. M. *The Effect of Resistance on the Intelligence Scores of Young Children.* (Teachers College, Columbia University, Child Development Monographs, No. 6, New York, 1931) 80 pp.

(84) RUST, M. M. "Factors influencing the validity of mental tests at the nursery-school age." *Proceedings* of the National Association for Nursery Education, Toronto, Ontario, 1933, 77–81.

(85) SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION. *The Intelligence of Scottish Children.* (University of London Press: London, 1933) 155 pp.

(86) SIMMONS, R. MCK. *A Study of a Group of Children of Exceptionally High Intelligence Quotient in Situations Partaking of the Nature of Suggestion.* (Dissertation, Teachers College, Columbia University, New York, 1938. To be published in Teachers College, Columbia University, Contributions to Education) 148 pp.

(87) STODDARD, G. D. "Child development: A new approach to education." *School and Soc.* 49: 1939, 33–38.

(88) STOKE, S. M., and LEHMAN, H. C. [The details of this reference unfortunately could not be supplied for inclusion here. — Ed.]

(89) STRONG, E. K., JR. *Vocational Aptitudes of Second-Generation Japanese in the United States.* (Stanford University Press: Stanford, Calif., 1933) 183 pp.

(90) TEAGARDEN, F. L. "Change of environment and the IQ." *Jour. Applied Psychol.*, 1927, 11, 289–296.

(91) TERMAN, L. M. *The Measurement of Intelligence: An Explanation of and a Complete Guide for the Use of the Stanford Revision and Extension of*

the Binet-Simon Intelligence Scale. (Houghton Mifflin Co.: Boston, 1916) 362 pp.

(92) TERMAN, L. M. *Genetic Studies of Genius. Vol I. Mental and Physical Traits of a Thousand Gifted Children.* (Stanford University Press: Stanford, Calif., 1925) 648 pp.

(93) THORNDIKE, E. L. "How may we improve the selection, training, and life-work of leaders?" *Teachers College Record,* 40: 1939, 593–605.

(94) THORNDIKE, E. L., and others. *The Measurement of Intelligence.* (Teachers College, Columbia University: New York, 1930) 616 pp.

(95) UNIVERSITY OF THE STATE OF NEW YORK. *Organizations of Special Classes for Subnormal Children.* (University of the State of New York Press: Albany, 1930) 9 pp.

(96) VINELAND LABORATORY. "Bibliographic notes: Publications on the Vineland Social Maturity Scale." *Training School Bull.,* 35: 1938, 57–59.

(97) VISHER, S. S. "The comparative rank of the American states." *Amer. Jour. Soc.,* 36: 1931, 735–757.

(98) VISHER, S. S. "Where our notables came from." *Scientific Mouthly,* 45: 1937, 172–177.

(99) WARREN, H. C. (Ed.) *Dictionary of Psychology.* (Houghton Mifflin: Boston, 1934) 371 pp.

(100) WELLMAN, B. L., and STODDARD, G. D. "The IQ: A problem in social construction." *Social Frontier* 5: 1939, 151–152.

(101) WELLMAN, B. L. "Our changing concept of intelligence." *Jour. Consult. Psychol.,* 2: 1938, 97–107.

(102) WHEELER, L. R. "A comparative study of the physical growth of dull children." *Jour. Educ. Research,* 20: 1929, 273–282.

(103) WITTY, P. A. *A Study of One Hundred Gifted Children.* (State Teachers College Studies in Education, Vol. 1, No. 13: Emporia, Kansas, 1930) 44 pp.

(104) WITTY, P. A. "The relative frequency of gifted boys and girls in the secondary school." *Educ. Admin. and Superv.,* 20: 1934, 606–612.

(105) WITTY, P. A., and JENKINS, M. D. "The case of 'B'—a gifted Negro girl." *Jour. Soc. Psychol.,* 6: 1935, 117–124.

(106) WITTY, P. A., and JENKINS, M. D. "Intra-race testing and Negro intelligence." *Jour. Psychol.,* 1: 1936, 179–192.

(107) WITTY, P. A. and LEHMAN, H. C. "The play behavior of fifty gifted children." *Jour. Educ. Psychol.,* 18: 1927, 259–265.

SECTION III

THE PHYSIOLOGY OF INTELLIGENCE

CHAPTER IV

THE PHYSIOLOGICAL CORRELATES OF INTELLIGENCE

LEONARD CARMICHAEL
President, Tufts College
Medford, Massachusetts
with special contributions by
MARY HARROWER ERICKSON, ROBERT C. TRYON, EDGAR A. DOLL, DONALD B. LINDSLEY, GEORGE KREEZER, JOHN R. KNOTT, and NATHAN W. SHOCK [1]

I. THE BODILY FUNCTIONS THAT ARE IMPORTANT DETERMINERS OF INTELLIGENT BEHAVIOR: INTRODUCTION

This chapter of the Yearbook, dealing with the bodily basis or correlates of intelligence, necessarily differs somewhat in scope from the preceding chapters. It is impossible to give, even in summary form, all the relevant evidence, which could be collected from a great variety of sources, that bears directly or indirectly upon the problem of the relationships of the make-up of the human organism and the correlated intelligence of the human individual. Among the various technical fields of knowledge in which information on this subject is to be found are anatomy, physiology, pathology, embryology, histology, anthropology, neurology, surgery, comparative psychology, and a dozen more subdivisions of these special fields. Keeping in view, therefore, the general purpose of this Yearbook, it seems best, in this chapter, first, to try to make clear the assumptions underlying the evidence that what we call ' intelligence ' is indeed grounded in the activity of the living organism, and second, to give some indication to the reader of the possibility of further study in this field, which is at once too technical and too extensive for full consideration here. Many of the bibliographical references given will be of especial value to the reader who wishes to go further into this subject and many of these books and articles themselves contain large lists of references.

[1] The portion of Chapter IV prepared by each of these special contributors is indicated in what follows. The remaining portions were all prepared by President Carmichael, to whom is due also the general organization of the chapter. — *Editor.*

II. Vitalistic Theories of Intelligence Not Subject to Scientific Study and Hence Not Treated Here

Intelligent acts, at least in man, have always been considered as related to the individual's so-called *mental life* as well as to the activities of his *body*. For many centuries, there has been an interest on the part of students concerned with mental phenomena in what is called the 'body-mind problem.' From the standpoint of the fundamental question of this chapter, this problem may be phrased as follows: What is the relationship between intelligent acts, including so-called 'conscious intelligent thought,' and an active brain and other active bodily organs? Today many students believe that the body-mind question is largely a verbal question, or what may be termed a 'pseudo-problem.' It is important, however, to notice presuppositions concerning the relationship between the organism and behavior that are today generally accepted by scientists, because from early times to the present, there have been certain purely mentalistic theories of intelligence and such theories cannot be assimilated into the general body of science. The philosopher Henri Bergson, for example, developed a theory of intuition and intelligence that is not subject to ordinary scientific investigation. The roots of all such vitalistic views, McDougall (128) has shown, trace themselves back into what is called 'primitive animism.' If one accepts such an animistic theory of the basis of intelligence, he may well not wish to continue reading this chapter, because a vitalistic view, if consistently held, essentially excludes the possibility of working out any *invariable* functional relationships between an intelligent act and any part of the body. The task of investigating this relationship is the aim of a scientific physiological psychology. If human intelligence can operate even occasionally without being related to the activity of the living body, the whole study of the nervous system, for example, in relation to behavior becomes speculation, not science. The undoubted facts disclosed by the brain surgeon, however, to take but one of the many possible lines of evidence, will keep most individuals from any serious consideration of such an essentially mystical theory of intelligence.

III. The Scientific Problem of the Relationship of the Living Organism and Intelligent Behavior

Careful observation, recording, and evaluation of behavior such as that involved in making appropriate marks on a printed intelligence test, in moving blocks in a formboard, or in opening a puzzle box are often first steps for one who would study the physiological psychology of intelligence. Then the student must attempt to discover how this specific behavior that is measured and compared with other known performances is correlated with the activity of certain parts of the body. Answers given by such work, although complicated and tentative, may be factual. The true relationships thus discovered must stand. Such facts will not be altered by any theorizing about the 'body-mind problem.' On the contrary, any such philosophical theory, to gain acceptance, must first demonstrate that it does not violate any of the known facts concerning the correlation of specific intelligent acts and specific brain functions that have been discovered by science. The formulation of theories in a field where there are no facts is not open to the same objection as is the formulation of theories that oppose known facts.

IV. The Problem of the Bodily 'Seat of Intelligence'

It is now clear that the total integrated organism is, in a way, always involved in any act that may be characterized as intelligent. For example, it is possible to show that some specific sagacious acts cannot be performed by an adult human being who has lost the big toe of his left foot that can be performed by one who has normal toes. But this does not mean that intelligence is located in the toes. In an understanding of intelligent acts, the brain has been determined by many years of differential experimentation to be centrally important. The sense organs, the whole nervous system, and the organs of reaction (the muscles and glands) are also involved in intelligent behavior (25). Over and above these organ systems themselves, however, must be considered the blood stream and its components, including oxygen, food materials, hormones, and the other agents that form the *internal environment* of the nervous system and, indeed, of the whole organism. This internal environment in definite ways conditions the growth and functional level of the brain and other organ systems, which themselves function in the determination of intelligent acts.

Given relatively normal sense organs and organs of reaction, it has been clear for centuries — although it was not always clear — that the brain, taken as a whole, is the most important part of the body in determining intelligent behavior. There is evidence that the Egyptian surgeons realized that head wounds specifically interfered with their patients' mental ability. Even the great Aristotle, however, confused certain of the brain's most important functions in intelligent behavior with those of the heart and described the brain as an organ involved in cooling the body.

Exactly what parts of the brain are exclusively important in intelligent acts is still of course a most complex problem, in which it is difficult to separate fact from speculation. Figure I shows a lateral

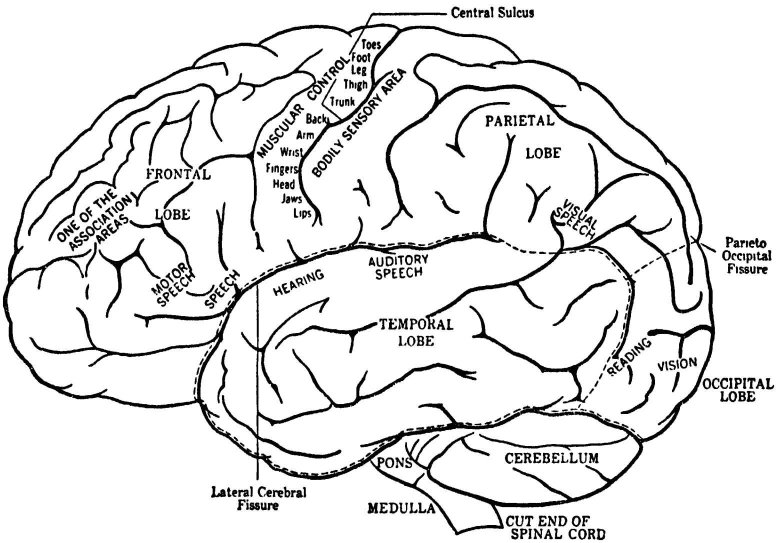

FIG. I.—SCHEMATIC DIAGRAM OF THE LEFT ASPECT OF THE CEREBRAL CORTEX

Some of the so-called 'motor' and 'sensory' areas are shown by labeling. Great caution must be used in interpreting such maps. Adapted from various sources. (See Carmichael, Ref. 25.)

view of the cerebral cortex with the principal lobes and centers indicated. In 1662 the great anatomist Stensen (51) characterized those who speculate about specific brain functions in the following terms:

> There abounds indeed a rich plenty of men to whom everything is clear. Such, dogmatizing with the utmost confidence, make up and publish the story of the brain and the use of its several parts with the same assuredness as if they had mastered with their actual eyes the structure of so admirable a machine and penetrated into the secrets of the great artificer. (p. 280)

Unfortunately, what Stensen said almost three hundred years ago is still true today (1).

V. Earlier Views on the 'Seat of Intelligence'

Historically the most notable of the misleading views concerning the localization of intelligence in the brain is that put forward by Gall in his system of phrenology (59). Gall, although in some ways a truly distinguished anatomist, worked out a theory concerning the relationship between intelligence and the brain, which, in spite of the efforts of modern proponents, such as Hollander (96), has definitely been disproved. (See Franz: 52.) Roughly, Gall's method consisted in observing casually and without measurement the psychological characteristics of various individuals. He also noted the skull shape of such living men. He also noted the skull shape of dead subjects whose brains he was able to dissect. He drew then the further conclusion that a depression on the skull corresponded with a depression in the brain and, still further, that a depression in the brain indicated a lack of brain function in this region. Therefore, he alleged that the region in question was the seat of the quality that he surmised, on the basis of his casual observations on living individuals, was lacking in the brain of the cadaver he was dissecting. In its extreme form, this theory never gained scientific support; rather, as a result of the investigations of such men as Bell (7), even during Gall's lifetime, it has always been taken as an example of scientific error. However, following the work of Gall, the pendulum of informed neurological opinion has several times swung back and forth from extreme views concerning the specificity of localization of functions in the brain. Marie (160), for example, successfully challenged some of the older views concerning great anatomical specificity in the brain of such centers as those involved in speech, but many other anatomists, such as Henschen (89), held that each particular cell of the brain had some specific *psychological,* as distinct from merely physiological, function to perform. The details of this controversy are well set forth in Piéron's work, *Thought and the Brain* (160) and need not be reviewed here. The reader should also consult Lashley (120) and Franz (54) in this connection. Little has happened in the last half century to challenge the words of Ferrier (44) when he wrote:

> It would, however, be absurd to speak of a special seat of intelligence or intellect in the brain. Intelligence and will have no local habitation distinct from the sensory and motor substrata of the cortex generally. There are centers for special forms of sensation and ideation, and centers for special motor activities and acquisitions, in response to and in asso-

ciation with the activity of sensory centers; and these in their respective cohesions, actions, and interactions form the substrata of mental operations in all their aspects and all their range. (p. 467)

VI. Specific Brain Mechanisms and Intelligent Behavior

It is possible for the neurologist to describe with some accuracy the major reflex centers of the brain below the cerebral hemispheres. The 'wiring diagram,' as it may be called, of the medulla, the pons, the thalamic region, and of many other subcortical centers has been worked out by a series of interesting techniques in a most satisfactory, although, of course, as yet incomplete, way. When we turn to the great cerebral hemispheres themselves, the anatomical problem becomes very much more complex.[1] The problem of structure and function here seems to yield best to a complex attack involving many techniques, including operative procedures on the nervous system with behavior study before and after, as well as a series of special embryological and physiological experimental procedures.[2] For example, excellent clues to the development of these so-called 'higher centers' are found in facts observed concerning the growth of the cerebral mechanisms in phylogeny; that is, in the evolution of the brain in the animal series. The olfactory portions of the cerebrum are almost as well developed in structural detail in the lower mammals as they are in man. It is far otherwise, however, with much of the rest of the cerebral hemispheres. For convenience in considering this evolution, the areas of the cerebral hemispheres, or specifically, 'the cortex,' have been divided into two parts, depending upon the phylogenetic history of the structures involved. The olfactory cortex, or archipallium (literally 'old mantle'), is thus termed the old brain and all the rest of the cortex, the neopallium ('new mantle'). This distinction is important because even in the simplest mammals the neuron systems below the cortex and the neuron systems of the archipallium have attained what may be very roughly characterized as their definitive form. But the mass of neuron systems are still very incompletely developed in the neocortex, or neopallium, of lower mammals.[3] Figure II shows in diagrammatic form the relative increase in the neocortex in the animal series.

In attempting to discover the bodily structures that make characteristic intelligent behavior possible, we must look with especial care at

[1] See Mettler (145) and Carmichael's tabular description of such methods (25).

[2] See Fulton (58) for a consideration of the physiology of the nervous system.

[3] See Kappers, Huber, and Crosby (107) and Bolton (13).

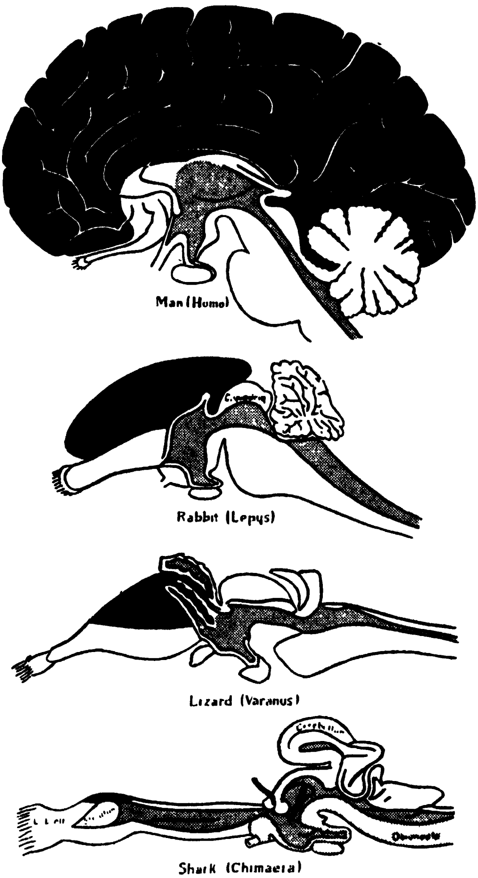

FIG. II. — A SCHEMATIC DRAWING OF THE 'OLD' AND THE 'NEW' BRAINS OF CERTAIN TYPICAL VERTEBRATES AND OF MAN

The striking relative increase in the mass of the new brain in passing from the shark (bottom) through the lizard and the rabbit to man (top) is shown in black. (From Edinger as presented in Koffka, Ref. 113.)

those parts of the cerebral hemispheres that constitute this neopallium. This later development is important because it is usually regarded as roughly correlative with the development of intelligent behavior. In this connection, we must remember that this new cortex may take over functions previously carried out by lower central nervous centers — or at least modify the more primitive functions of the lower centers. This process of taking over functions by the cortex has been termed the 'corticalization of function' as a continuation of the more general process of 'encephalization'; that is, of the development of the dominance of central nervous ganglia in the head region (39). Roughly speaking, this corticalization of function is seen as the brain develops in the animal series and in the growth of the human brain during prenatal life and in childhood (13, 36).

It is generally agreed that in the process of the development of the great new cortex, or neopallium, sense organs, such as the eyes and ears, have played an important rôle. Organisms that are most adequately equipped with special distance receptors are best able to meet changing environmental problems and hence to survive. Behavior acts that result from stimuli the immediate sources of which are removed in time and space from the reacting organisms are of course identical in many instances with those we characterize as intelligent. Hence, it is most significant that there is an observed parallelism in the development of distance receptors and of the brain mechanisms that make such distance receptors effective, and the organism's ability to make those reactions we term intelligent. It is interesting to note how often 'insight' depends on 'sight,' for example.[1]

VII. Learning Ability and Intelligence

Learning ability is closely associated with the capacity to perform intelligent acts of the sort just described. For this and for other reasons, it is not surprising that we see the qualitative improvements in learning ability often considered as closely associated with the extent of development of the new cortex. This is not to be taken as meaning that the anatomical, histological, or other direct study of the brain, in absence of known facts concerning the individual's life achievement, makes it possible, save in extreme cases of defect, to say anything concerning the 'mental qualities' of an individual (172). We must remember, however, that even in organisms with a central nervous system made up of a chain of ganglia, some profiting by past experience is possible, as is shown in the experiments of Yerkes (209) and others on habit functions in the earthworm. It is even possible, as many authors have contended and as Coghill (26) specifically has recently said, that "any protoplasm can learn"; that is, any protoplasm, in a primitive way, can behave 'intelligently.' Surely, however, in the human being, that protoplasm which is specialized to make

[1] For a more complete discussion of this important field, see Munn (152) and the references given by him.

plasticity of behavior possible is primarily localized in the brain and in that part of it that we have here called the 'neopallium.' Concerning this new brain, Sherrington has recently said with real eloquence: "Before it, truly, there were educable systems in the animal world, but this is so educable as to be practically a new thing in the world" (182). Recent experiments by Culler and his associates (71) and others have demonstrated that in animals from which the cortex has been removed simple associative learning of the conditioned response type is still possible. Brown and his associates (68) have also shown the part of subcortical mechanisms in learning. Thus, the modifiability of behavior essential to very simple intelligent acts can, at least in a most primitive way, be determined by the lower brain centers in the higher mammals.[1] In this connection the large number of adaptive acts carried out by cats after the removal of the cortex is most significant for an analysis of behavior, as has been shown by Bard and Rioch (5).

Experimental study has made it necessary to ask whether the relationship between intelligent behavior and the cerebral cortex is the same at all levels of the animal series, no matter what the stage of development may be of the neopallium itself. That is, can we generalize from facts concerning intelligent behavior and the brain of the rat, cat, or chimpanzee to man, or must the relationship between the brain mechanism and intelligence be worked out separately for each animal form?

Different answers have been given to this question by competent modern students (206). For the purpose of this chapter, however, it seems best to say that we should not apply *directly* to man any knowledge concerning the relation of brain mechanism to intelligence gained from the study of the lower mammals or even from the anthropoid apes. We must recognize, nevertheless, that such knowledge is valuable in its own right and also exceedingly suggestive concerning relations that are apt later to be discovered in the case of man. For example, Lashley's large series of most important and careful studies has not disclosed any specific locus of association centers, as such, in the rat's brain that may be considered as basic to learning or intelligence (120, 121, 122, 123). Rather, all his investigations demonstrate that, save for specific cortical sensory projection areas, such as the visual centers, learning and intelligence in the rat are a function of the *amount* of cortex present that functions, as he says,

[1] See, on this point, Frolov (57).

equipotentially; that is, the mass of cortex present, quantitatively considered, rather than any particular region, is basic to intelligent behavior in the rat. In this connection the experiments of Maier (136, 137, 138), Krechevsky (114), Cameron (18), Ericksen (42), Melton (144), and others are important.

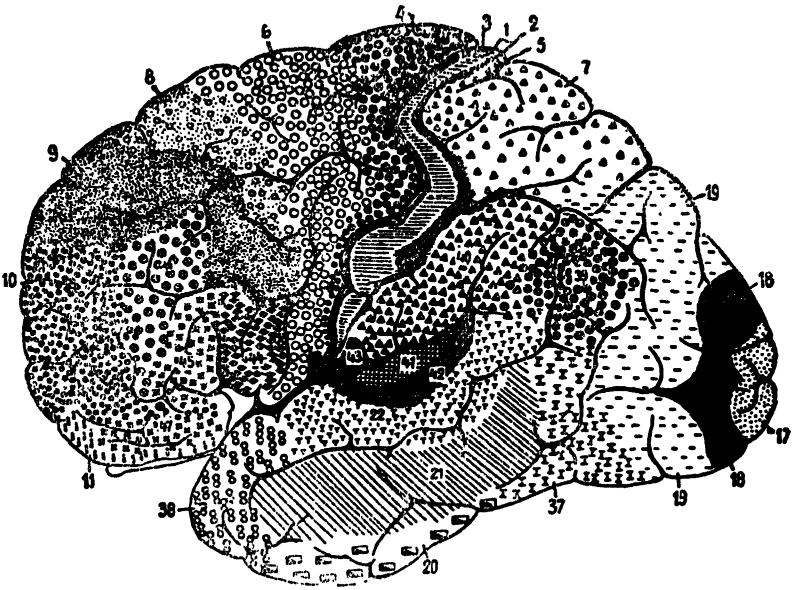

FIG. III.—THE SPECIALIZED ANATOMICAL OR 'CYTOARCHITECHTONIC FIELDS' OF THE LATERAL ASPECT OF THE HUMAN BRAIN

Each symbol stands for a well-differentiated cell field. In all, 50 or more such fields are differentiated. (For details, see Herrick, Ref. 90, and the other references noted by him.)

Some recent studies of human brains injured as a result of war wounds, tumors, surgery, or localized pathology have tended in a general way to confirm many of Lashley's findings on the rat. This does not mean, of course, that the years of intensive study of the human brain that have made it possible to map with great assurance the more detailed sensory and motor regions of the cortex, in relation to topographical region and cell character (see Figure III), are in error. All this intensive work has not led to any universally agreed upon centers that are to be regarded as the specific seats of general intelligence or of learning ability. (See Figure I.) This is true in general in spite of the fact that a number of different regions in the brain have been shown in specific clinical cases to be related to specific forms of intelligent acts, such as those used in language and other types of symbolism, although under proper motivation, remarkable achievements in re-education are possible, as demonstrated by Franz (52, 53, 55).

Recently also the effect of cerebral destruction on so-called 'instinctive,' as contrasted with individual-adaptive or intelligent, responses has been studied with results not unlike the earlier studies on learned reactions.[1] In this connection it must be remembered that a spread of functional disorganization following any sort of wound, of the sort termed 'diaschesis' by von Monakow (151), may obscure individual clinical or experimental findings as Pike has indicated (161). Or, as Henry Head puts it, in his famous chapter on *Language and Thinking* in his work on aphasia (87), the loss of intelligence in aphasic patients (that is, patients with a language disability attributable to the central nervous system) may be very specific. He says: "When some act or process is disturbed in consequence of an organic or functional lesion, the abnormal manifestations are the result of fresh integrations carried out by all available portions of the nervous system" (p. 532). Nevertheless, as Pike and Chappell (162) point out, physiologically "there is little evidence that there is . . . an 'equipotentiality of the cortex.' . . . In any organization, certain parts have certain functions." This does not mean that a mass of cells may not have mass functions, but that when there is anatomical specificity, there may well also be some specificity of function.

VIII. Clinical Evidence on Intelligence and the Brain

The best clinical view now available seems to be that much brain activity basic to the human behavior we term 'synthetic intellectual activity' takes place in the temporal, occipital, parietal, and frontal lobes of both hemispheres of the cerebrum. Dandy (28), a distinguished brain surgeon, goes further than this and asserts that *since:* (1) the right cerebral hemisphere above the basal ganglia can be removed with no pronounced disturbance of 'mentality,' (2) the same can be said after removal of both frontal lobes, (3) the same can be said after the removal of much of the occipital lobes, and (4) the same can be said after the removal of the lower third of the temporal lobes, *therefore* intellect is normally related to parts of the parietal lobe and other lobes especially involved in the cortical speech mechanisms of the left cerebral hemisphere. He also points out that cutting off the blood

[1] See, for example, Stone (189) and Lashley (121).

supply of the anterior cerebral artery on the left side *always* leads to complete 'loss of consciousness,' whereas cutting off the blood supply of this same artery on the right side does not lead to that outcome. His observations also lead him to believe that the corpus callosum, the great connecting system between hemispheres, subserves no general mental function.

The frontal lobes, however, do play an important though obscure part in determining intelligent behavior. The loss of the frontal lobes of the brain does not change simple visual-auditory associations that have already been established or the ability to make such associations. This slender clue has led to a most intensive consideration of the rôle of the frontal lobes in the determination of the higher forms of animal and human intelligent behavior. The present evidence seems to indicate that a loss of the frontal lobes rather subtly affects the entire mental life of the human individual. The emotional and motivational, as well as the intellectual, behavior is changed characteristically, though sometimes only very subtly, in unfortunate individuals who have lost large parts of their frontal lobes. This observed fact suggests that the frontal lobes are involved in the intellectual functions characterized by 'drive.' The frontal lobes also seem to involve the capacity of an individual to synthesize past learned acts or associations into very complex conceptual wholes. For a review of this whole field, see Mettler (146). Brickner (16), a recent student in this field, has said on the basis of clinical evidence:

> The frontal lobes add to the intricacy of the intellect by receiving material from all the rest of the cortex and by making complex associated aggregations of it, but there is no evidence that they serve as centers for the intellect in general or for any particular intellectual function. (p. 302)

Continuing this point of view, Brickner contends that the frontal lobes may be thought of as cell masses that make possible the 'enrichment' of intelligent behavior by virtue of the large multiplication of associative possibilities their presence makes possible. Thus, the frontal lobes are sometimes conceived of as biological luxuries, and half in fun they have been described as parasitic growths by such evolutionists as the late W. M. Wheeler.

IX. Further Consideration of the Frontal Lobes and Intelligence

A review of a large number of cases involving the removal of all, or of parts of, the frontal lobes leads to many statements such as: "After the operation, the patient found it more difficult to do mental arithmetic and play bridge than before." "The patient showed loss of initiative." "The patient showed a lack of capacity for planned administration." Certain of these studies, notably one by German and Fox (62), suggest that the change in intelligent behavior determined as a result of brain operations bears a relationship to specific topographical cerebral localization and to the dominance of the right or left hemisphere, rather than a change in behavior related to sheer cerebral mass, as a strictly interpreted theory of equipotentiality would suggest. The facts of so-called 'dominance' of one hemisphere over the other, in relation to the visual and speech mechanisms of man, must always be remembered in interpreting clinical studies.

In one of the most carefully controlled studies of frontal-lobe function, an intelligent patient was studied intensively for many months (15). Because of a tumor, a large part of both frontal lobes had to be removed (the right specimen weighed 108 grams, the left specimen, 121 grams, thus making a total cerebral loss of 229 grams). This operation was performed by Dandy. The subject was a stockbroker who owned a seat on the New York Stock Exchange. As a result of the operation just described, the patient's personality underwent certain changes, but they were less intellectual than emotional alterations. Whereas before the operation the patient had been rather taciturn, he now became boastful. After the operation, he showed less regard for the feelings and welfare of his family than before. The simple elements of old and new associative material, however, could still be completely understood by the subject. He seemed less able, however, to utilize diverse material in forming complex syntheses than had been characteristic of him before. Yet he was still able to solve mathematical problems involving the use of algebra, and it was still possible for him to memorize a poem when he was asked to do so.

This case is interesting in relation to the famous 'crow-bar case' (85, 11) reported many years ago in which a laborer lost a large part of the frontal part of the brain in a blasting accident and showed very little intellectual change following his recovery from the accident though his temper became more violent and his general emotional control less effective than it had been. Recent operations on the frontal lobes in an effort to cure certain types of psychosis that involve emotional disturbances are interesting in view of the facts just presented.

Section X of this chapter, dealing with the relation in general between brain lesion and mental function, was prepared by a student of experimental psychology active in this field, Dr. Mary Harrower Erickson, who is associated with the well-known brain surgeon, Dr. Wilder Penfield, in the Montreal Neurological Institute of McGill University. One who wishes to study this field in detail should also consult the important series of papers upon the experimental study of the frontal lobes in primates by Jacobsen (101, 102, 103).

X. Brain Lesions and Mental Functions

M. R. Harrower Erickson
Montreal Neurological Institute
McGill University
Montreal, Canada

A review of the literature concerning the relationship of localized lesions of the brain to the concomitant mental phenomena [1] is startling in its lack of uniformity. One finds reports of cases in which marked mental changes are ascribed to definite lesions, side by side with those in which lack of all disturbances, or virtually no disturbance, is noted.

Dandy (28), for instance, reports no mental changes after removal of the left occipital, left temporal lobes, and even after removal of the right hemisphere. Rowe (174) cites a case of a patient who, also after removal of the right hemisphere, despite " relatively mild mental changes " is still able to perform very creditably on the Stanford-Binet Intelligence Test. Alford (1) claims that there is no intellectual defect produced by any lesion except one in the central portion of the left hemisphere in right-handed persons, and a small area of the optic thalamus. Jefferson (106) finds in frontal lobe removals that " those who showed no mental alterations before operation were unaffected by partial removal of the anatomical frontal lobe " and that " those with mental symptoms were much better after the lobe had been excised." Hebb and Penfield (to be published shortly) find none of the so-called ' frontal lobe signs ' in a case of bilateral frontal lobectomy. The patient, who had been unmanageable and uncontrollable before operation for removal of scar tissue, became a normal member of his family after (IQ 98). The presence of pathological tissue prior to operation in this case, and its absence afterwards, may afford a valuable clue to the understanding of the relation of a lesion to mental functions.

---

[1] No mention is made here of specific sensory or motor disturbances (86), or of the aphasias (207). Both disorders would require separate treatment.

On the other hand we have evidence of various types of intellectual disturbances and personality changes associated with widely varying areas. In the frontal lobes we have Goldstein's extensive studies (76, 77, 78, 79, 80, 81, 82) emphasizing the loss of abstract attitude and a return to the more concrete level of thought and behavior: "a disintegration of a higher, more complicated process, determined by the whole organism, to a less complicated lower one, determined by the stimuli of the outer world." There is Brickner's (16) exhaustive report of a case of bilateral frontal lobectomy with unmistakable intellectual deterioration, vulgarity, and personality changes. There are large series of cases, such as Frazier's (56), where it is found that one or more symptoms characteristic of mental disturbance occurs in 60 of the 105 cases of frontal lobe tumor studied. Loss of memory, loss of intelligence, lack of attention, change in personality are among the symptoms listed. Strauss and Keschner (192) found that there were disturbances of "intellect and the higher psychic functions" in 62 percent of their group of 85 cases of frontal lobe tumors — disturbances manifested chiefly by a "diminution or loss of the capacity to synthesize simple thought processes into more complex thinking." Among others who have investigated and commented on changes of mentality in frontal lobe cases are Piotrowski (164), Sachs (175, 176), Schwab (179), Williamson (208), Henry (88), Penfield and Evans (157), and Halstead (study to be published shortly).

Similarly, in the case of the temporal lobe Keschner, Bender, and Strauss (112) found changes in "intellect and higher psychic functions" in 62 patients (56 percent of the cases studied). They consider that, since logical thinking and correct reasoning "depend on the ability to synthesize concepts and impressions, . . . a pathologic process . . . affecting the anatomic or physiologic substratum for perception, memory, and association must necessarily interfere with proper synthesis." They find no difference in the nature of these temporal-lobe symptoms and those obtained with the group of frontal-lobe patients. Kennedy (111), German and Fox (62), Gibbs (69), and Goldstein, among others, have discussed the disturbances arising from lesions in this area.

In addition to these regions most commonly associated with psychological functions, specific mention has been made of psychological changes occurring when lesions involve the parietal lobe, hypothalamus, posterior fossa, and corpus callosum.

Various factors have been cited as contributory causes for these mental changes. For example, the rapid rise of intercranial pressure is held by Southerland (187), Strauss and Keschner (192), Holmes (97), and Pessin (158) to be of great importance. The extent of brain tissue involved and the rapidity of the growth are mentioned by Strauss and Keschner also. The toxic effects from tumor tissue is held as an important factor by some writers, but is denied by others (Jefferson, 106).

To those who are perhaps over-anxious to ascribe a definite psychological function to a specific area, an important note of warning is sounded in the numerous writings of Goldstein (78–82). For example, "Circumscribed foci of the cortex never lead [exceptions indicated] to psychic changes confined to one province, but always all performances are more or less changed." And again, "abnormal phenomena manifested by brain patients are only in small part . . . the direct result of the loss of the injured place; for the most part they are the expression of the changed function of the new psychological situation which has been created by the defect." Again, "The fundamental disturbance of different patients show an essential affinity independent of the locus of the lesion."

In support of this, experimental work now in progress (by the writer) indicates that a recognizable picture of a restricted, limited, and devitalized personality, as estimated by the Rorschach test (see Piotrowski, Ref. 164) is found in the cases of infiltrating tumor, pre- and postoperatively, regardless of the location of the lesion. The writer has also found, however, that in cases of clean surgical removal of scar tissue or atrophied brain, a somewhat different picture is presented. Preoperatively these patients show records that are similar to those of the preoperative tumor group, but postoperatively an expansion of psychic reactivity, not found in the tumor group, may occur. This does not mean that the postoperative records are always normal, for marked individual personality difficulties appear in almost every case, but the typical, uniformly restricted picture presented pre- and postoperatively by the infiltrating tumor group is not found. Since in this latter case we are dealing, as nearly as possible, with the mere absence of tissue, whereas in the infiltrating tumor group it may be assumed that pathological tissue is still present, these results suggest that some factors other than, or in addition to, the location of the lesion must play a part in determining the nature of the mental changes, when they occur concomitantly with a given lesion. It should be noted, however, that not all methods of estimating changed mentality are equally valuable. For example, a patient studied by the writer, after removal of the right frontal lobe for an infiltrating tumor, achieved an IQ of 135 on the Stanford-Binet, but showed on the Rorschach test a personality unmistakably altered by her pathological condition.

Several factors, therefore, may account for the apparent contradictions in the literature concerning the question of mental changes concomitant with cerebral lesions: first, the use of different tests by different authors (some tests are unquestionably more suited to reflect alterations in mentality and behavior than others); second, the fact that the criterion of 'altered mentality' as used by different investigators is by no means uniform; third, the fact that cases which are, on the face of it, identical for comparative purposes may actually involve very different types of pathological processes, even assuming that the location is identical.

XI. The Contribution of the Study of the Mentally Deficient to an Understanding of the Physiological Basis of Intelligence

Mentally deficient human beings (or aments) and those who are mentally ill (or dements) have been studied for years because of the importance of such individuals to society, whose charge they are. The fact that such individuals are often institutionalized has made their study easier. Also, as stigmatized individuals, it has been possible to investigate familial resemblances and the like as is not always possible with the more normal members of the community.[1] Moreover, today few seriously attempt to support the view that aments, geniuses, and normal individuals are distinct human species (24). Rather, all levels of human ability are generally conceded as falling into a quantitative continuum regarding the possession of all measurable traits. Thus, it appears very likely that facts discovered concerning the transmission of feeblemindedness will have direct bearing on the familial transmission of other levels of ability, although strictly speaking, this fact is still to be demonstrated in man (74).

It may safely be said that current medical teaching is overwhelmingly that represented by Tredgold (197), when he designates as a cause of what is termed 'primary amentia' an hereditary factor dependent upon germinal variations that have their locus in the chromosomes.[2] Primary amentia was defined by Tredgold as

> a state of restricted potentiality for, or arrest of, cerebral development, in consequence of which the person affected is incapable at maturity of so adapting himself to his environment or to the requirements of the community as to maintain existence independently of supervision or external support. (pp. 8–9) [3]

Amentia of this type, therefore, is held to be inherited, although the exact mechanism of this inheritance may differ from individual to individual and may be correlated with varying constitutional abnormalities. Intelligence may not in any sense be thought of as a simple genetic trait. Students of feeblemindedness are, therefore, interested in the relations between feeblemindedness and many environmental and

---

[1] See Town's recent study *Familial Feeble-Mindedness* (196).

[2] See Sherlock (180).

[3] In the 1937 edition of the book, this definition has been amplified by Tredgold.

somatic characteristics that may be demonstrated to have an hereditary basis. Alcoholism, syphilis, tuberculosis, the age of the parents at the time of the child's conception, injury of the fetus during pregnancy by mechanical or toxic means, abnormal labor, infantile brain wounds, and the like are all at times related to amentia. The exact part that heredity and environment play in cases of feeblemindedness that seem to be 'caused' by such factors is often very difficult, if not impossible, to determine. Each of the topics just enumerated has its own large bibliography. But in evaluating the evidence, for example, of the causation of feeblemindedness by the alcoholism of the parents, great care must be taken to exclude genetic selective factors, and unfortunately in most clinical studies, such factors have not been, and probably cannot be, excluded. The effect of treatment upon the measured IQ of patients is very variable in general paresis (syphilis of the nervous system). In one case with an IQ of 72, before treatment, a change to 102 two months after treatment is recorded, but the opposite effect has also been noted (41).

Among the more important historical studies on the inheritance of mental ability may be mentioned Goddard (72, 73, 74, 75) and Dugdale (38).

XII. The Contribution of Specific Genetic Experiments to the Question of the Inheritance of Intelligence

Case studies of defective human stocks and investigations of royal pedigrees still leave much to be desired in giving a definite answer to the problem of the inheritance of mental ability. Studies of human twins have provided the best direct human evidence on the so-called 'nature-nurture' problem.[1] For this reason psychologists have been forced to turn to lower animals where controlled conditions may be maintained in order to investigate this important field. The work of Heron (90) deserves special consideration, as does that of Tryon (see 198–203 and his special contribution under XIII, farther on).

In a series of experiments Heron has demonstrated that certain abilities in maze-running in rats are subject to genetic selection. In his work, techniques have been used that are essentially the same as those used to determine the inheritance of hair pigmentation or eye-color in rodents in the now well-established independent science of genetics.

---

[1] For the report of a most significant study in this field and bibliographical references, see Newman and Freeman (153).

In this work, however, one must be careful not to assume that excellent performance on the maze necessarily measures the same processes that are measured by human intelligence tests, though it seems clear that maze-running is in one sense a test of the differential adaptive capacity of the organism and hence a type of behavior relevant to the general problem under consideration.

In the material presented below Tryon has summarized his own most important experiments in this field:

XIII. Genetic Differences in Maze-Learning Ability in Rats

Robert Choate Tryon
Associate Professor of Psychology
University of California
Berkeley, California

The experimental geneticist is inclined to be somewhat skeptical of conclusions from nature-nurture studies on human beings. The complex breeding system deriving from the adventitious circumstances of 'young love' and the wide differences in environments into which the progeny of such matings are born and reared present such a complex matrix of determiners of 'mental' differences that it would appear hopeless to endeavor to separate out the relative effects of those termed 'nature' and those termed 'nurture.' Furthermore, the indubitable tendency in the uncontrolled human environment for different genotypes to seek out different environments, and conversely, for different environments to select special genotypes, and all this to an unknown degree, creates a correlation between nature and nurture that confounds any analyst who has neither an hereditary nor an environmental axe to grind. And on top of it all, to make the confusion worse, are the ambiguous psychological omnibus tests of mental ability, about the psychological validity of which the most able psychologists cannot agree.

After several centuries of the type of groping called 'natural history,' the biologist in his rôle as experimental geneticist has discovered that the only way to get definite answers to the nature-nurture question in plants and animals is to establish a pure strain experimentally by means of a controlled selective breeding schedule and then experimentally to vary systematically the milieu for different samples of the strain. The logic is, indeed, very simple: In a given species, (1) hold heredity constant by choosing a pure strain, then study the effects of different environment on it, and (2) hold the environment constant, then vary heredity by studying the development of different pure strains in it, and (3) compare the relative effects of the two types of variations that are, respectively, nature and nurture. Though the logic is simple, its experimental execution is arduous, often requiring many years of work even when investigating only one character in one species.

As the psychologist cannot, of course, perform such experiments on human beings, he must turn to animals. During the last three decades, the animal psychologist has developed techniques for the reliable and valid measurement of individual differences in numerous psychological characters of animals, especially of rats. The securing of definite answers to the question of the relative effects of nature and nurture on psychological characters is thus made possible. An experimental genetics focused on animal *behavior* and ultimately based on studies of numerous species of animals varying in phylogenetic complexity should give us the answers we wish. Only the most egregious 'special creationist' would argue that the findings of such a comparative psychological genetics would have no applications to man.

The attempts of the writer to design and execute such a psychological genetics experiment are outlined below. The essential aims have been to establish under environmental control a maze-bright and a maze-dull strain of rats, to determine the nature of the genetic determiners at work, to discover the constancy of this psychological difference throughout a large range of the rats' life span, and to find important biological and psychological correlates of the differences in this maze ability. Finally, the effects of systematic environmental changes on each strain are to be investigated. Though the work has now been in progress eleven years, it is in many respects still preliminary. Findings are complete with respect to some of the objectives but not to others. A brief summary is presented below under the various types of analyses.

1. Proof of the Inheritance of Individual Differences in Maze Ability

An experiment [1] was begun in 1927 that had as its purpose the establishment by selective breeding of a pure line of maze-bright and a pure line of maze-dull rats. Each animal was run nineteen trials through a seventeen-blind T maze. His score was the total number of entrances into blind alleys. The breeding schedule consisted in mating together the brightest rats within each of the brightest litters, the dullest within each of the dullest. Rigorous environmental controls were effected (1) by instituting standard procedure of animal care and of breeding, (2) by using an automatic mechanical device for delivering the animals into the maze without handling, and (3) by employing an electric recorder for the scoring of each rat's maze run. These controls have remained constant for eleven years. Selective breeding has been

[1] Supported by grants from the National Research Council, the Carnegie Corporation of New York, the Research Board of the University of California. Aid in statistical analysis was provided by the Works Progress Administration under Official Project No. 465-03-3-631-A2 at the University of California.

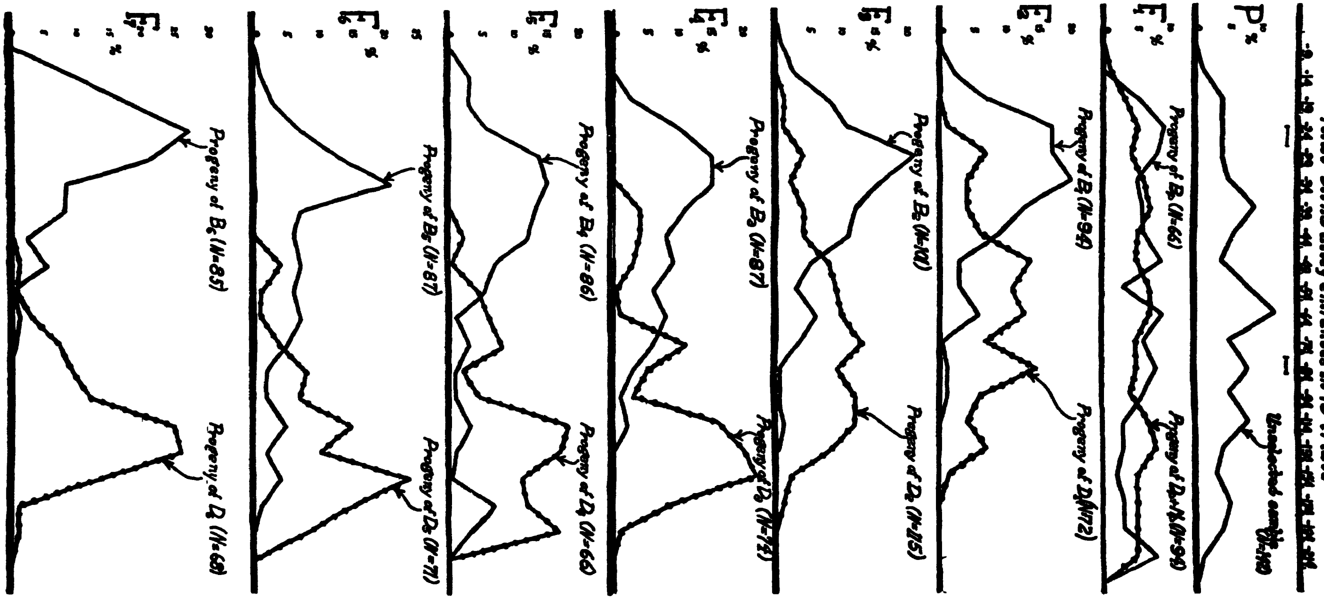

FIG. IV. — EFFECTS OF SELECTIVE BREEDING ON MAZE-LEARNING

Along the top is the scale of brightness as evidenced by the total number of blind-alley entrances made in nineteen trials. All the distributions below use this common top scale. For instance, a bright animal who made from ten to fourteen errors would fall under the scale step, the upper limit of which is marked —14, a dull who made from 195 to 214 errors would fall under —214, etc. The first generation of rats, marked "P" to the left, is shown just below the scale. The total number of P rats was 142, and the percent of them lying at each point on the scale is indicated in the distribution. The brightest of these were bred together, and then the dullest, giving the two F, groups, as shown. The selective breeding effects are shown down to the F_7, where progeny of B_6 (bright F_6) are markedly different from the progeny of D_6 (dull F_6).

continued for eighteen generations. As success in establishing strains of bright and dull animals is crucial to the proposed project, I present in Figure IV the basic data showing the effects of selective breeding through the F_7 generation. For brevity, I have not presented all the later generations, but

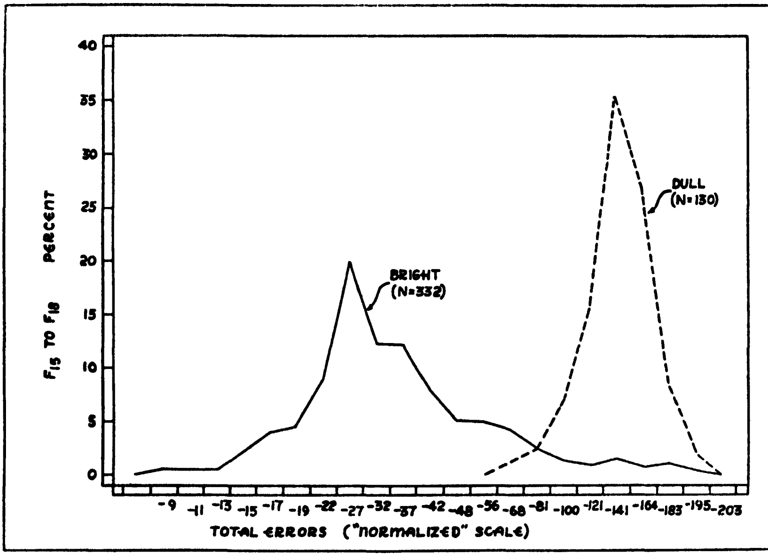

FIG. V.—BRIGHT AND DULL STRAINS OF THE FIFTEENTH TO EIGHTEENTH GENERATIONS

to show the latest results, I have given in Figure V the results in the F_{15}, F_{16}, F_{17}, and F_{18} generations. In this figure the distributions of the two strains are shown for these later generations combined. There appears to be a law of diminishing returns, for after the F_7 negligible effects of selective breeding are noted. The results for all generations will be depicted in final form in terms of the improved normalized scale.

2. The Genetic Basis of Differences in Maze-Learning Ability

What is the genetic factor basis of differences in learning ability? How many factors must be postulated and what is the nature of their interaction? Geneticists propose a multiple cumulative-factor theory as the genetic explanation of the plant and animal characters that have statistical frequencies similar to those of maze-learning. One crucial experimental test of this theory is the cross between the pure lines at the two extremes of the scale. The F_1 progeny of such a cross should show a homogeneous median performance. The next generation progeny of the F_1 should vary widely over the whole scale. Figure VI shows the actual results of such a test on our behavior

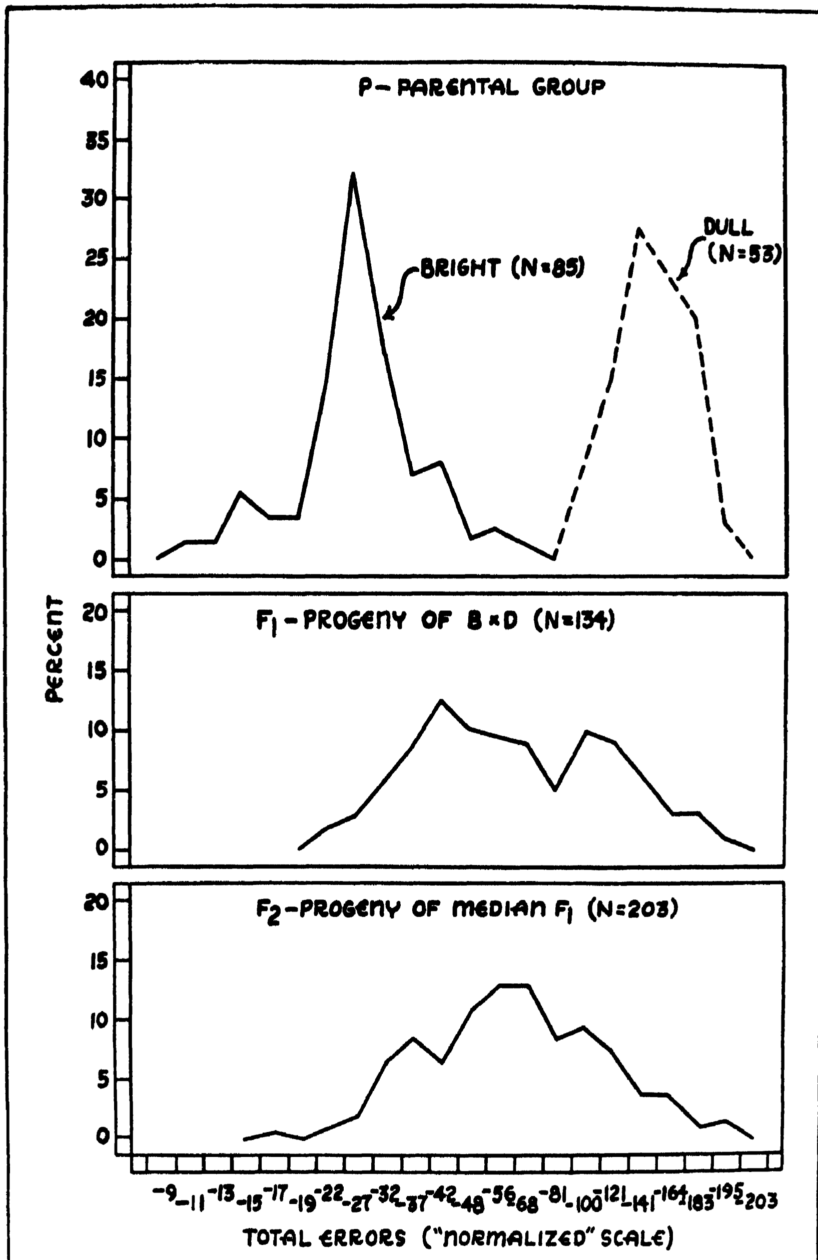

Fig. VI. — F_1 and F_2 of Bright × Dull trait. The F_2 progeny of the F_1 do *not* vary more than the F_1. Because the going theory was not verified, I repeated this crucial type of cross several years after the first attempt, but the results were the same. Figure VI combines the findings of both series of experiments.

We need to develop and verify a factor hypothesis consistent with the facts of Figures IV, V, and VI, and with the results of certain back-cross experiments. One method is that of expressing a given hypothesis concretely in terms of a dice-pattern. A random parental population is then set up in which the score of each fictitious individual is determined by a dice throw. These 'individuals' are then 'bred' according to the actual breeding schedule of the rat experiment and the theoretical results compared for fit with the experimental findings. The hypothesis that gives the best fit and that is most consistent with genetic theory will be the one finally chosen. I have performed several such artificial experiments, and from these it appears that the most promising hypothesis investigated to date is one that postulates multiple factors, some dominant for bright performance, some (but fewer) dominant for dull, and some cumulative. The effects of linkage and crossing-over, and of reciprocal crossing — that is, bright male × dull female vs. dull male × bright female — must be investigated.

3. Biological Correlates of Brightness and Dullness

Large groups of bright and dull animals have been carefully measured in respect to brain size and weight, body weight, and fertility. Significant differences have been discovered. The bright animals show physical superiority throughout, except in the one particular of fertility.[1] Because of the intensive inbreeding that has occurred during the process of selective breeding, the evaluation of these findings is to some extent equivocal.

4. The Constancy of Differences in Ability

The extent to which brightness and dullness persist throughout the lives of the animals was investigated. One hundred and seven animals were measured in maze ability when they were young; then an interval of seven months, amounting to about two-thirds of a rat's life span, was permitted to elapse before remeasurement. The correlation between the early and late measurements was .80, indicating a high degree of constancy. For different subgroups that experienced different degrees of environmental variation during the intervening period between measurements, the correlation was the same — a result suggesting that environment as varied in these experiments plays a negligible rôle in this ability.

---

[1] The difference between the N's of the bright and dull groups of Figures IV, V, and VI is a function of the number of matings made by the experimenter and not of differential fertility.

5. Psychological Nature of Brightness and Dullness

A variety of experiments were conducted that aimed to provide some insight into the psychological nature of the differences between the bright and dull groups.

a. Evidence from Ratings of Emotional Characteristics. Using reliable, objective rating scales, carefully standardized on a large group of preëxperimental animals, three judges independently rated 234 bright, dull, and stock animals on their hiding, avoidance, and escape reactions to controlled handling by the experimenter, and on their reactions to novel inanimate objects in the maze situation. The results show clearly that the bright animals are most adjusted 'emotionally' in the maze-learning situation, whereas in response to handling they are 'neurotic.' Exactly the reverse is the case for dulls.

b. Evidence from a Cluster Analysis. In addition to the measures of emotionality, this same group of 234 animals was measured on eleven different aspects of efficiency at different stages of learning the maze; for example, efficiency as measured by errors, speed (rate) of running, hesitation time at choice-points. The significant result from the analysis appears to be that the hereditary difference in maze ability, which was itself discovered by selection on the basis of errors only, is reflected in all the measures of efficiency as well as in those of emotionality.

c. Evidence from Experiments on the Sensory Nature of Ability Differences. Are bright animals superior to dull because of superiority in *sense acuity,* or does their superiority reside in a capacity for *abstracting* the spatial relations of the complex maze path? On the sense acuity hypothesis, one would assume that during the 19 trials in the maze the bright animals had learned to follow visual, auditory, kinesthetic, tactual, and olfactory cues by virtue of genetic superiority in their senses. Five experiments were conducted to investigate this hypothesis. In each experiment about 70 bright and 70 dull animals were subjects. Stated briefly, the technique was that of experimentally disrupting cues of the various sense modalities on the trials following the nineteenth. In every experiment the bright animals showed relatively negligible disturbance; many showed no disturbance at all as a consequence of cue disruption. These results fail to support the sense theory, but rather support the view that bright animals are superior to dull in a capacity to *generalize* the spatial pattern.

d. Evidence from an Analysis of the 'Qualitative' Behavior of Bright and Dull Animals in the Maze Situation. To investigate further the differences between bright and dull animals in the nonsensory determinants of their maze behavior, I analyzed the frequency pattern of errors made in the 17 blind alleys of the maze by 500 bright and 500 dull animals. As the error patterns of the two types of animals were quite different, especially in the later stages of learning, I attempted to deduce the existence of a number of psychological gradients of a nonsensory spatial character that theoretically determined the behavior of brights and dulls. As the postulation of such gradients constituted being 'wise after the event,' I then attempted to predict the pattern of errors made by 150 animals in a quite different twenty-blind maze. The correlation between the predicted error pattern and the actual error frequencies was .70, and with minor changes in the weights of the gradients, the correlation was .92. The hypothesis of nonsensory gradients is therefore supported.

e. Evidence from the Study of the Behavior of Maze-Bright and Maze-Dull Animals in Other Problem-Solving Situations. To what degree does hereditary brightness and dullness represent a *general* capacity to learn? In one experiment 150 animals were run through another maze. In another study, conducted by Krechevsky, bright and dull groups were studied in a brightness discrimination box. At the present time Searle is engaged on a program of observing the performance of bright and dull animals in a number of tasks requiring discrimination of distance, angles, and brightness. From the work done and under way, and from published data of other workers showing the relation between maze-learning and other types of learning, the evidence clearly supports the view that maze-learning is *specific*. The doctrine of alleged 'general ability' supported by many psychologists has not been substantiated in rats.

6. Prospectus

Future lines of research are clearly indicated.

One of these, as pointed out earlier, is the study of the effects of experimentally induced environmental variations on the two strains. The general question is this: What sorts of environmental variables of a psychological and biologically pathological character will make hereditarily bright animals dull, and hereditarily dull animals bright? No systematic experiments of this sort have yet been performed. The only relevant data we have are from the experiments on constancy and cue variations cited above. After the rats had first learned the maze to the best of their capacities, the introduction of unsystematic though extensive variations in milieu in different groups during the interval between the original learning and retest did not affect the test-retest correlations. In the stimulus disruption tests, extraordinary variations or deprivations of stimulus (that is, environmental) features did not significantly affect the bright-dull differences. But it is to be noted that these environmental variations occurred *after* the maze had been learned to the level of capacity. These observations would seem therefore to verify the hypothesis that after hereditary capacity has expressed itself in the building up of habits and concepts, the effects of gross environmental changes are negligible. On the other hand, we know as yet nothing about the effects of gross milieu changes before and during learning.

Another problem that calls for study is a systematic examination of the physical growth and behavior development of the *young* of the bright and dull lines. An extensive series of observations was made on the sensory and motor development of 50 progeny of bright and 50 of dull from birth to 30 days of age. These observations were purely exploratory. A program of research in this field, based on this preliminary work, is to be instituted.[1]

In Section XIV Dr. Edgar A. Doll summarizes the present state of knowledge concerning the psychological effects of cerebral birth lesions on the intellectual achievements of patients so afflicted. The reader should also consult in this connection the excellently worked-out case of Gesell (63, 67).

XIV. Psychological Consequences of Cerebral Birth Lesions

Edgar A. Doll
Director of Research, The Training School
Vineland, New Jersey

Congenital palsy is probably as old as the race. Modern attention to this problem usually dates from 1839, when William J. Little, an English orthopedist, described some of the consequences of birth lesions. The physical condition has since generally been termed 'Little's disease.' More recently the term 'birth injuries' has been employed as a popular expression for this condition in its more general aspects. This is a somewhat unfortunate term, since to many it connotes obstetrical trauma, whereas, as a matter of fact, congenital birth lesions may occur at almost any stage of the reproductive process from late pregnancy to early neonatal life. Prenatal infection, trauma, malformations, or other pathology may be present before birth. Prematurity of birth, with weakness in the circulatory system producing brain hemorrhage, is a further consideration. In the birth process itself pelvic malformations, primaparity, type of presentation, size of the infant, and other considerations may induce complications of labor. Labor may be prolonged, severe, assisted by instruments, modified by drugs, and so forth. These abnormal conditions are some of the hazards of delivery that obstetrical skill may not be able to surmount.

Likewise, deficient animation at birth, ill-advised methods of resuscitation, early postnatal trauma from the handling of the infant, neonatal infectious diseases, and other pathology producing essentially the same effects as obstetrical or prenatal embarrassments are commonly grouped among the gross complications that may attend the reproductive process as a whole.

[1] References to previous publications on the experiments described in Section XIII are 198, 199, 200, 201, 202, 203, 204, and 205.

For this reason the term 'intracranial birth lesions' without reference to specific etiology is preferred to the term 'birth injuries,' but the latter already has definite hold on terminology. Birth lesions are not restricted to intracranial lesions, though they are most commonly and most seriously of this type. One must also reckon with developmental anomalies of the central nervous system.

Such brain damage is generally rather severe, but many minor lesions, such as petechial hemorrhages, may go unnoticed because of their relatively slight consequences. There is good reason to believe that such minor damage is far more prevalent than is commonly supposed, being reflected in minor incoördination, lack of motor grace, mildly unstable personality, and mild reductions from the hereditary potentials for intelligence.

Birth lesions most conspicuously produce motor impairment, especially spastic paralysis and athetosis. These impairments are most evident in the face, trunk, and limbs, but may involve any muscle groups; for example, those of visual coördination, breathing, and the less obvious motor functions of the entire organism.

Impairment of the intellectual system is often either overestimated or underestimated because of the more obvious motor symptoms. Intellectual impairment may produce any degree of retardation, even to profound idiocy, with or without motor accompaniment. Or the intelligence may be unaffected. The majority of birth-injured subjects are of normal intelligence, and many are of superior intelligence. Among these, one can only estimate how much higher the intelligence might have been without brain damage.

The percentage of birth lesions in the general population is not accurately known, but it is probably not more than one per thousand of births surviving the first year of life. Among birth-injured subjects as a whole, about one-third are mentally deficient, which is from ten to twenty times the percentage of the mentally deficient in the general population. Among the feebleminded, about 10 percent owe their condition to birth lesions. According to present figures this is the most frequent cause of mental deficiency other than heredity.

Another consequence of birth injury is impairment of personality. This phase of the subject is less clearly understood, but competent authorities attribute both gross and minor disturbances of behavior and disposition to this presumptive cause. This may be produced directly by way of damage to the autonomic nervous system, or may be indirect, due to the frustrations of behavior imposed by the motor condition.

There is a further presumption that birth injuries may produce sensory impairment either by way of the motor accompaniments of sensation, as in muscle sense, balance, and visual coördination, or by damage to cerebral areas that subserve sensation and their associated areas. Likewise, speech may be affected not only through its motor components, but also through the cerebral preconditions of language.

Some of the functional impairment resulting from cerebral birth lesions may be offset in the course of time by organic repair or by the transfer of function from one part of the brain to another. Thus, one hemisphere may to a degree take over the functions of the other, or these functions may in time or as a result of treatment be relegated to the lower nerve centers. In general the damage is permanent and nonprogressive. Delayed, as opposed to arrested, mental development is frequently encountered.

The social competence of the individual is modified by the nature and extent of such functional impairments. Since intelligence, motor facility, personality, language, habits, skills, interests, and the like, are determining factors in social attainment, it is obvious that their impairment as a result of birth lesions will be reflected in social competence. Indeed, the measurement of social competence, as for example by means of the Vineland Social Maturity Scale, is especially important in tracing the course of development and the influence of training or therapy in such patients.

The consequences of birth lesions are specially important in relation to education. They limit the opportunities and methods of education in each patient according to the direction and extent of the handicap. If intelligence is impaired without motor handicap, the person so affected may be treated as simply mentally deficient. As the motor effects are more or less severe, the teacher must reckon with the particular physical abnormalities that may be reflected in such fundamental activities as speech, reading, and writing.

The measurement of intelligence in such cases is particularly difficult, since that measurement is dependent upon motor expression of some sort. This calls for special insight, skill, and experience if the examiner is to avoid the dangers both of overestimation and of underestimation. The range of psychometric procedures is definitely limited by the nature and severity of the handicaps that may be present, and the tests employed must be selected and administered with special caution. Indeed, the examiner is frequently put to the test of psychological observation without benefit of standard psychometric procedure, and must rely upon his own judgment of the subject's comprehension of range of ideas, information, memory, judgment, attention span, and the traditional faculties of psychology. The examiner must also reckon with the environmental limitations imposed by the handicaps. The environment may be apparently rich but really poor, since the patient may be unable to capitalize the opportunities that surround him. Patience, tact, perseverance, and imagination are necessary for adapting methods and content of instruction to the limitations reflected by the handicaps.

It is especially important to note that the majority of birth-injured children are not mentally deficient and that a considerable number of these are of better than average intelligence. As noted previously, some of these patients have notably superior intelligence in spite of extreme motor handicaps, and this fact is of special significance. The intellectual maturation of these normal children apparently proceeds at a normal rate in spite of the extreme limitations to expression. The development of language in such cases, even when articulate speech is totally lacking, is of particular psychological import. The biographies of mentally superior patients are especially illuminating, as for example that of Hoopes (100). The histories of such individuals are educationally most illuminating and afford encouragement to both teachers and therapists to whose attention such patients may come.[1]

XV. Brain Weight and Intelligence

Somewhat varying average weights for normal adult human brains have been given in the literature. It is suggested that there are racial differences in brain weight, but this fact has not been demonstrated without question (119). The mean weight of the normal adult male brain may perhaps be taken as about 1,374 grams, but male brains weighing as little as 1,100 grams have been possessed by 'normal' individuals. The mean weight of the adult female brain may be set at 1,294 grams. A brain of a female idiot weighing only 227 grams is on record, as shown in Bolton (13).[2] There are great variations in the weights of the brains of individuals who show in every way a normal mental life. The brain of Cuvier, the naturalist, for example, weighed as much as 1,830 grams, but the brain of Housmann, the mineralogist, weighed only 1,226 grams.[3]

Because of the striking variations in brain weight, however, as well as because of the increase in the relative weight of the brain to the body, which seems to characterize the evolution of more and more intelligent animals, there has been a long history of an attempted scientific correlation between brain weight and intelligence. In all such work, the relation of brain weight to body weight, not absolute brain weight, is considered. In some such work, the brain itself is not too surely divided into two parts, the so-called 'somatoencephalon' and the 'psychencephalon' (14, 12); that is, the parts of the brain that are supposed to be related to lower reflex activities of the organism, and the parts related to the adjustment of the organism to its environment or its 'mental functions,' respectively. Anthony (3) has summarized the

[1] See also References 33, 34, 38, 40, 45, 50, 127, 134, 159, 173, 191.

[2] The individual who possessed this brain showed no signs of ordinary human mental life but lived a so-called vegetative existence. The brain stem and cerebellum were disproportionately large. It is estimated that the development of this brain in essential characteristics was arrested in the fourth fetal month. This brain in some ways resembled the brains of certain of the lower mammals but it is best thought of as an incompletely developed human brain rather than as an "evolutionary reversion to a prehuman type."

[3] For other brain weights see Bastian (6).

literature on brain weight and intelligence, and more recently, Lapicque (119) has discussed this same material in detail. At the present time, the work of Dubois (37) is most often taken as the starting point for discussions of this topic. In the older work in this field, an illegitimate effort to guess at the degree of intelligence of different animals studied often vitiated the conclusions drawn. Today, under the influence of von Bonin (14), the problem has shifted to a consideration of the possibility of devising a generalized mathematical statement concerning the relation between body weight and brain weight without reference to assumed intelligence. Once the facts of body-brain weight have been worked out in a series of animals, the question arises whether the study of animal behavior may show any correlated relationship with intelligence. Von Bonin accepts the formula $E = kS^r$, in which E is taken as brain weight in grams, k as a constant, S as body weight in grams, and r as .655. The constant, k (the so-called 'cephalization coefficient'), is determined from species to species. When k is thus taken as a means of classifying animals, that of man is found to be .99, or practically unity, and that of certain of the primates as high as .71, that of certain of the carnivora, .61, and that of certain of the rodents, .16. The relationships between these determined constants, like the failure to find any new and significant anatomical basis for human superiority in mental ability over the animals, is most significant. This finding points to the conclusion — also otherwise well substantiated — that man's intellectual processes that are qualitatively in a class by themselves are dependent upon, it may be, a relatively slight improvement *quantitatively* in the ability to use symbols and language. Thus a slight improvement of a certain sort in the brain gives man his overwhelming advantage of being able to profit by the past experience of the race as contained in concepts and symbols.

Efforts have also been made to work out a generalized description of the relation between brain and body weight in men and women separately and then to study the relation between a constant in such equations and estimated individual differences in mental ability. But save for some not over-sure efforts to claim racial superiority on the basis of such data, no review of this material is indicated here. Paterson has given an excellent review of the errors of popularized knowledge in this and related fields (155).

That the tremendous growth in brain weight in infancy and childhood is not without its significance in the differences in intelligent behavior between children and adults goes without saying. The average male brain may be taken as weighing about 330 grams at birth, 916 grams at one year, 990 at ten years, and 1,374 grams at twenty years or above (36, 197).[1]

XVI. The Growth of the Individual Brain in Relation to Intelligent Action

Studies of recent years on prematurely delivered infants and late human fetuses indicate that behavior of a complex and apparently adaptive sort is possible in the prenatal period (22, 24, 98). It has not been demonstrated, however, that the modifiability of behavior, which is characteristic of most specific intelligent acts, is developed, at least in any marked way, before birth. This statement is made in spite of the fact that Ray (171) and Spelt (188) have secured some indication of conditioning in the unborn human infant. The work of Marquis (140) and of Kasatkin and Levikova (108, 109), however, demonstrates that the first true conditioned responses cannot apparently be formed until some time after birth. It is true that the difference in the internal environment during the latter part of pregnancy and during the first few days of postnatal life may account for many differences in functional capacity. Students of the neonatal period (first two weeks of postnatal life) are agreed that almost all action of which the child is then capable is of the sort characterized as 'inborn' rather than of the sort resulting from specific individual 'learning.'[2] There is, however, some statistical indication that the absolute zero of intelligence should be placed in the prenatal period. The logic of this suggestion concerning the extension of the mental growth curve beyond its empirical date, we cannot go into here.

It should be remembered that at the time of birth, the average brain weight of the human infant is about one quarter the weight of an average adult brain. Thus in the course of childhood there is a large opportunity for development of total brain mass. In this development an elaborate series of histological and so-called 'cyto-architectonic' changes take place. There still seems to be no evidence, however, to contradict the old statement of Donaldson (36) that after birth no *new* brain cells are produced as a result of cell division, and the loss of brain cells in adult life is permanent and not subject to true regeneration. On the other hand, a great deal of activity and growth in the brain cells already laid down undoubtedly occurs during

[1] For a consideration of the problems of bodily growth see Scammon (178).

[2] See Pratt (167) and Peiper (156).

each individual's life. The exact importance during this time of the development of the myelin sheaths of effective neuron mechanisms is still a debated question (2, 107). Tilney (194, 195) in recent years advocated the view that the functional development was absolutely correlated with myelinization. This position was first brought into prominence by the work of Flechsig (47, 48), whose principles governing the periods of myelinization of the cortex are illustrated in Figure VII.

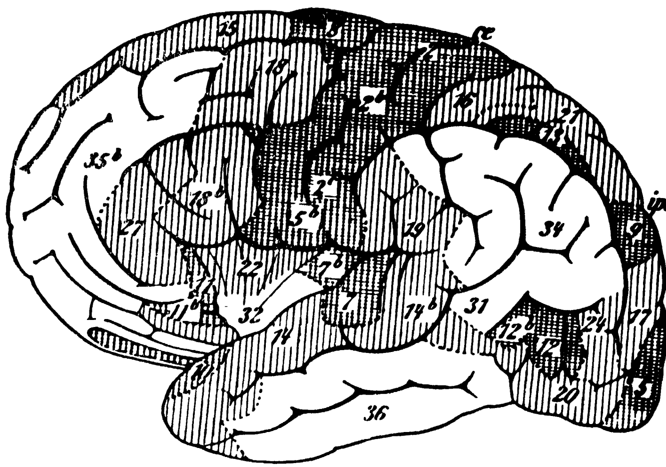

FIG. VII.—LATERAL ASPECT OF THE BRAIN SHOWING DIFFERENT REGIONS IN RELATION TO THE TIME OF MYELINIZATION

The regions first myelinated and presumably first in functional maturity are in dark cross-hatching, those next to be myelinated, in lighter cross-hatching, and the last, in white. (From Ranson's, 171, diagram of Flechsig.) Numbers represent approximately the order in which different parts secure mature fibers. See also Herrick (90).

It would be going far beyond the facts to allege that the growth of intelligent behavior that is measured during the first, second, third, fourth, and subsequent years can be thought of as a mere correlate of the changes resulting from the inner *growth* of the nervous system during this period. That much behavior change during this period is dependent upon sheer growth rather than upon specific environmental change, however, is suggested by analogy with the functional changes that accompany the growth of the lower brain centers during fetal life. Minkowski (149), for example, has worked out most ingeniously a story of the gradual 'encephalization' of function that involves the control of the plantar reflex; that is, the reflex that involves movement of the foot, and especially of the toes, following stimulation of the sole of the feet (148).[1] This sequence begins with a local, non-neural, purely muscular response and progresses by stages, once neural control is established, from the spinal level to final cortical control. All these alterations in behavior are directly dependent upon a specific form of growth of the central nervous system that makes changes in behavior resulting from the same sort of stimulation possible.[2]

There is thus good evidence that alterations in the character of intelligent behavior during early years are related to the gradual maturation (that is, growth in all its details) of the brain. In all this development, the exact part that may be played by 'exercise' (external environmental stimulation) is still open to debate. It seems clear, however, that exercise at one time does not have the same effect that it will have at a later time. This fact is demonstrated by a number of studies, of which those of McGraw (133), Kellogg (110), and Jacobsen, Jacobsen, and Yoshioka (102a) are typical. Evidence seems to be piling up, moreover, that specific function plays a negligible part in the growth and differentiation of the cerebral hemispheres, as Burr (17) has recently shown on the basis of studies in experimental embryology. This is not in any sense a refutation of the established facts concerning neurobiotaxis (107); that is, the fact that growth in the nervous system is related to the stimulation of non-active cells by active ones.

XVII. The General Problem of the Relation of the Electrical Phenomena of the Brain to Behavior

The study of the electrical responses of the human brain has recently offered an important new tool to investigate the activity of the normal intact human nervous system. In the three subsections of this chapter that follow, three experimental students of this field summarize the relation between the so-called electroencephalogram findings and the development of ability, and between them and different levels of ability. My own acquaintance with the comparative aspects of this field leads me to believe that it may be one of great significance in the future, once the techniques now being actively improved have

[1] See also Dewey (31).

[2] In this connection the analogy of such a simple developmental sequence and that worked out by Gesell and his collaborators for the maturation of complex adaptive acts in the infant is most attractive (65, 66, 67). See also Hooker's analysis of grasping behavior (99).

been fully developed. The present status of the field in relation to the topic of this chapter is well given in the next three sections. The reader will note that there is still a difference of opinion concerning the relation of the recorded electrical phenomena of the brain and the aspects of behavior that are called 'intelligent,' but it must be remembered that this technique, which in its magnifying power is comparable to the microscope or the telescope, is as yet very new. The three contributions prepared especially for this portion of the Yearbook by Dr. D. B. Lindsley, Dr. George Kreezer, and Dr. John R. Knott follow without additional comment.

XVIII. The Ontogenetic Development of Brain Potentials in Human Subjects

Donald B. Lindsley
Director, Psychological Clinic and Neurophysiological Laboratory
The Bradley Hospital and Brown University
Providence, Rhode Island

It is now well known that minute electrical potentials can be recorded from the surface of the head in intact human subjects and are indicative of an underlying activity of the brain. The graphic record of the electrical activity of the brain, known as the electroencephalogram (abbreviated EEG), was first described for human subjects by Berger (8) in 1929. Characteristic of the EEG as described by Berger were two types of rhythmic waves known as 'alpha waves' and 'beta waves.' Both these rhythms are present to a lesser or a greater extent over all regions of the head in normal adult subjects, but the alpha rhythm is much larger and more prominent in most regions, and for that reason has proved more amenable to study. During the ten years since Berger's first publication on the subject many studies by workers in this country and abroad have described additional characteristics of the EEG as recorded from different regions of the head under a variety of conditions. The most recent and extensive reviews of the subject are those by Jasper (104) and Davis (29).

Although analysis of the EEG records has revealed other types of waves, the frequency and amplitude of alpha waves and the percent of time an alpha rhythm is present constitute the main types of quantitative analysis applied in the study of such records. In normal adult subjects most investigators agree that the approximately 5 to 100 microvolt alpha wave potentials have an average frequency of about 10 per second with a range of variation in different subjects from 8 to 13 per second. The percentage of time that such a rhythm is present varies from almost zero to 100 in different subjects and depends, as do the other characteristics of the rhythm, upon the region of the head from which the records are obtained.

The first evidence that the electrical activity of the cortex had an ontogenetic history of significance came from Berger (10), who discovered that the alpha waves of the EEG first appeared a few months after birth and that the frequency of the waves when present in young children was lower than in adults. Loomis, Harvey, and Hobart (126) and Davis and Davis (30) also found lower alpha wave frequencies in children than in adults. In a study of 54 adults and more than 100 children Lindsley (124) found that series of rhythmic alpha waves from the occipital region of the head first appeared in infants at three to four months of age. From an initial frequency of 3 to 4 per second the waves increased with age until the adult average of 10 per second was reached by eight to ten years of age. Attention was called to the fact that the time of onset of the alpha waves in the occipital region corresponds roughly to developing visual perceptive functions, particularly the ability to follow a moving object across the visual field.

Smith, in a study of 65 children from one day to seventeen years of age, confirmed Lindsley's results and added important observations on the electrical activity of the sensory-motor region, which appeared considerably earlier than that in the occipital region. More recent studies by Lindsley (125, 125a) and Smith (184, 185), involving larger numbers of children of all ages from birth to maturity, and utilizing both cross-sectional and longitudinal methods, have further confirmed the earlier results with respect to the time of onset of the occipital alpha rhythm and are in close agreement with respect to the growth in frequency of the waves with age. The frequency of the occipital alpha waves increases quite rapidly during the first two years from an initial value of 3 to 4 per second at the time of onset to about 6 per second at the end of the first year, and 7 to 8 per second by the end of the second year. Thereafter the increase is less rapid, reaching the lower limit of the adult range at five or six years of age and the adult average at between eight and twelve years of age. Smith (185) called attention to the exponential trend of the curve of increasing frequency, and Lindsley (125) stressed its similarity to the curve of brain growth (brain weight). The latter investigator found also that four other types of waves (including beta waves) that he was able to identify in his records increased in frequency according to the same general function that described the increase in the frequency of the alpha waves with age.

Smith's (185) observations on the electrical activity of the sensory-motor region indicate that there may be three or four independent rhythms, ranging in frequency from 1 to 20 per second (1 to 3, 3½ to 6, 7 to 8, 12 to 14, and sometimes 18 to 19), present in that region at or within a few days after birth, but only in infants when asleep or in the process of going to sleep. His first observations suggested that these rhythms changed relatively little with increasing age. More recently he has published a report (186) showing that one of them, the 7 to 8 per second rhythm, increases in frequency after one year of age in much the same manner as the occipital alpha rhythm.

In addition to the frequency of the waves, other factors, such as regularity, amplitude, and the percent of time that a recognizable alpha rhythm is present, have been studied in children. Both the sensory-motor and occipital alpha rhythms show an increase in the amplitude and regularity of the waves, as well as an increase in the percent of time that they are present in a given record during the first year or so while the rhythms are becoming established, but Lindsley (125) has found that after the second or third year the amplitude declines gradually to a somewhat lower average adult level as do also the percent time and regularity.

Kreezer (115, 116), on the basis of EEG's from subjects with various types of mental deficiency where chronological age could be, in effect, held constant (adult subjects with low mental ages) and mental age could be used as the variable, has found some evidence that frequency, amplitude, and percent time of alpha rhythm are related to mental age rather than to chronological age. Other evidence has failed to bear this out, however, and the significance of his results must be questioned, pending more complete presentation of the data with evidence of its statistical significance. The EEG records of four mentally deficient subjects (IQ's ranging from 43 to 74) studied by Lindsley (125) showed no significant differences in frequency, pattern, or percent time of alpha rhythm from those of normal children of the same chronological age. Furthermore, the curve describing the increase in frequency of the alpha waves with age in the 256 children studied by him was distinctly different from that which described the increase in mental age. Also, there was no significant relationship between IQ and the characteristics of the EEG studied within any age level.

In summary it may be said that the ontogenetic development of cortical potentials apparently begins before birth — this is in agreement with findings by Jasper, Bridgman, and Carmichael (105) in the guinea pig — since there is evidence of various rhythms over the sensory-motor region at birth. Only one of them, however, the 7 to 8 per second rhythm, which later probably becomes the 'precentral' alpha rhythm, shows evidence of growth in frequency and that only after the first year of age. The presence of these sensory-motor rhythms at birth or earlier is in agreement with our knowledge of the functional state of these cortical areas as evidenced by behavioral responses. The appearance of the rhythms in the sensory-motor region earlier than in other areas is also in agreement with the anatomical-histological evidence of earlier development and maturation of the structures in this region. The time of onset of the occipital alpha rhythm (about three to four months of age) suggests a rough correlation with the functional development of the visual cortex as shown by behavioral indices, particularly with regard to postural and prehensory adjustments involving visual perception. Although the developmental trend of the alpha frequencies appears to be more closely identified with processes of physical growth and maturation of the brain, it is perhaps not unreasonable to suggest that some characteristic of the EEG may prove to be related in a significant way to some aspect of mental growth. So far this has not been clearly and unequivocally demonstrated, but the ontogenetic study of cortical potentials is only just begun and will undoubtedly provide in the future further important information with regard to the functional and structural development of the cortex.

XIX. The Relation of Intelligence Level and the Electroencephalogram

George Kreezer
Assistant Professor of Psychology
Cornell University
Ithaca, New York

Since many important properties of a tissue are lost when the tissue dies, it has seemed of importance, in the search for a neural correlate of intelligence, to find methods that would make possible examination of activity in the living brain. The electroencephalographic technique makes possible the recording of the electric activity of the brain by means of electrodes attached to the external scalp, and without need for direct exposure of the cortex.[1]

The method of recording the electroencephalogram (EEG) is, in principle, simple. Very minute fluctuations of electric potential (about the order of 50 millionths of a volt) are picked up by electrodes attached to the scalp, increased in magnitude by means of a vacuum-tube amplifier, and finally recorded as a fluctuating line by means of a photographic or ink-writing oscillograph. The fluctuations of potential may be classified as the alpha waves, the beta waves, and the delta waves. The chief criterion used in distinguishing these different types of wave is that of frequency or duration. Thus, in the normal adult, alpha waves show an average frequency of about 10 a second; beta waves, of about 25 a second; and delta waves, a duration, for individual waves, greater than $\frac{1}{8}$ of a second. In a series of investigations conducted by Kreezer at the Vineland Training School and Letchworth Village[2] on the relation of the EEG to intelligence level among the mentally deficient, the properties of the EEG that have been considered are alpha index, alpha wave frequency, average alpha wave amplitude, and delta index. Alpha index represents the percentage of the total time that rhythmic sequences of alpha waves are present in a standard record. The delta index, defined in a some-

[1] Further details concerning the electroencephalographic technique and reviews of recent investigations may be found in Jasper (104) and Kreezer (119).

[2] Detailed results of most of the studies summarized here have not yet been published. Preliminary reports may be found in References 115, 116, and 117.

what different way, is a measure of the prevalence of delta waves in a given record. Alpha wave frequency refers to the average number of alpha waves per second in different sequences of alpha waves, and alpha amplitude to the average amplitude of alpha waves throughout a standard sample.

In the investigation of the relation of intelligence level to these properties of the EEG, it has been necessary to control a number of variables that might influence the EEG independently of intelligence level. These have included types of mental deficiency, chronological age, sex, emotional excitability, inter-electrode resistance, electrode location, and conditions of the subject during the recording. Since the experimental groups were limited to subjects above sixteen years of age chronologically, mental age (Stanford-Binet) could be used directly as an index of intelligence level.

The relation of mental age to the electroencephalogram was investigated separately for three different types of mental deficiency.

First, in a group of 50 subjects of the *mongolian type,* ranging in mental age from 1.5 to 7.5 years, the following product-moment correlations were obtained: between mental age and alpha index, +.348, between mental age and alpha amplitude +.313, and between mental age and alpha frequency −.213 (118a). The first two correlations are statistically reliable in terms of the criteria proposed by Fisher (46); the correlation with alpha frequency is not. To test the correctness of these results, a second series of mongols was investigated. In a series of forty-eight subjects so far examined, a significant correlation of mental-age level and alpha index (+.324) was again obtained, but non-significant correlations for alpha frequency, and delta index.

Second, in a group of 46 subjects of the *nondifferentiated familial type* of mental deficiency, no significant relation was found between Binet mental age and alpha index, alpha amplitude, or delta index. In the case of alpha *frequency,* however, a value for the correlation coefficient just on the borderline of statistical significance was obtained ($r = +.32$). In a second series of subjects of this type being investigated, there is again evidence of a positive correlation between alpha frequency and Binet mental age, but the magnitude and reliability of the correlation has not been determined because the series is not yet sufficiently extensive.

Third, in a group of 13 subjects with *phenyl pyruvic amentia,* a statistically significant correlation of +.72 was obtained between Binet mental age and alpha index, but no significant correlation of mental age with alpha frequency or delta index. This correlation coefficient should probably be taken as tentative, in view of the relatively small number of subjects in the group.[1]

In addition to the foregoing studies, reports upon the relations of the EEG to intelligence level have been made by Rahm and Williams, and by Lindsley.

---

[1] Since subjects of this type occur with a frequency of only one in about every 25,000 of the general population, it is relatively difficult to obtain a sufficiently large series. Further investigations are in progress.

Although differences in the conditions of these studies make comparison with the investigations already considered somewhat difficult, the results tend to be confirmatory.

The experimental group of Rahm and Williams (169) was made up of 36 unselected mentally deficient individuals ranging in mental age from under 1 year to 9.8 years. Since the group contained only two Mongols, it seems likely that it is to be regarded as more closely comparable with the nondifferentiated familial group considered above than to either the Mongolian or phenyl pyruvic group. The correlation obtained between percent time alpha [1] (alpha index) and mental age was +.20. This coefficient was increased to +.27 when corrected for variations in life age. Neither correlation is statistically significant, a result in agreement with the nonsignificant correlation of +.16 found by Kreezer for the nondifferentiated familial group.

In a group of mentally normal individuals ranging from 10 to 14 years of age chronologically, Lindsley (125) obtained a correlation coefficient of +.02 (obviously nonsignificant statistically) between intelligence-test score and alpha index. This result, too, is probably most directly comparable with the results for the nondifferentiated familial type of mental deficiency, since the latter type is commonly regarded as making up the lower region of the normal distribution curve of the general population.

These results described suggest the following conclusions:

1. Among the mentally deficient, there is evidence of significant correlations between properties of the EEG and mental-age level. These findings suggest that the EEG will probably be of value as an index of neurological factors associated with intelligence.[2] They do not, however, justify its use as an objective substitute for intelligence tests. The relations obtained hold for groups of subjects and not for individuals.

2. The fact that not the same property of the EEG was found to correlate significantly with mental age in different types of mental deficiency indicates that the physiological factors associated with intelligence are multiple, and that different ones of these factors are affected in different types of mental deficiency.

3. The finding, in the non-differentiated familial group, that alpha frequency alone was significantly correlated with mental age is of interest in connection with the reports that alpha frequency increases with life age in normal children (125). This correspondence suggests that the developmental change in frequency found in normal children may be associated with mental age rather than life age as such.

4. The most important need in further research is the extention of the in-

[1] Certain differences occur in the criteria used by different investigators in calculating alpha index, or 'percent time alpha.'

[2] Preliminary physiological hypotheses have been proposed concerning the nature of the cortical conditions responsible for the properties of the EEG found to vary significantly with mental age level (115, 116).

vestigation to sufficiently large and controlled groups of subjects to permit completely decisive answers to be made to the questions raised. Related physiological investigations will then rest upon a sounder basis.

XX. Comments upon Some Electroencephalographic Correlates of Intelligence

John R. Knott
Research Associate in Psychology
State University of Iowa
Iowa City, Iowa

The recently developed techniques for recording minute changes of electrical potential in the brain are obviously inviting to those who seek after the cortical mechanisms underlying intelligence.

The record of the electrical activity of the cortex is best envisaged as being composed of a continuum of frequencies, ranging from very slow (0.1 to 1.0 cycles per second) to very fast (50, or more, cycles per second), and varying in amplitude. In many individuals, when the record is secured under the so-called 'standard conditions' — eyes closed, room dark, subject mentally and physically relaxed — there is a predominance of rhythmic activity at a frequency of about ten cycles, with another predominance at a frequency of about twenty cycles. The rhythms of these frequency bands have been labelled *alpha* and *beta,* respectively. The alpha rhythm is most prominent over the visual area, while the beta rhythm is most prominent over the motor area. Most records in most investigations have been taken from the visual area.

Since the alpha rhythm seems to be that most clearly affected by sensory stimulation, 'attention,' and the like, and since there is a progressive increase in its frequency from 3 months to 10 years of age, the efforts of most research workers have been directed toward this one area of the whole continuum.

The developmental aspect of the alpha rhythm of the EEG has served as an opening wedge in the investigation of the electrocortical correlates of intelligence, for this increase in alpha frequency could be a function of mental age or of chronological age. The first observations bearing on this issue were made by Berger (9), the discoverer of the human EEG. He noted that a normal frequency of the alpha rhythm was observed in several feebleminded patients, that it was slowed in idiots who had never learned to speak, and that in all these cases the amplitude appeared below the average of normal subjects.

More recently, a comprehensive study of the EEG in relation to intelligence has been undertaken by Kreezer. His technique has been to study individuals of relatively high chronological age but of relatively low mental age; this rules out the influence of chronological age *per se* on the EEG. Kreezer's measurements of electrocortical activity included not only alpha frequency and alpha amplitude, but also a measure of the percentage of time that regular sequences of rhythmic alpha waves appeared.

In his first report, Kreezer (115) observed that patients presenting the mongolian and the hereditary types of mental deficiency showed all the electrocortical rhythms usually observed in normal adults, but that within the clinical groups there was a smaller proportion of regular sequences of alpha waves in the records of subjects below MA 5 than there was in subjects above that level. In these same subjects, he observed large, slow, 5-per-second waves, particularly over the motor and frontal regions; these waves were not observed at mental age levels over five years. Normal children between 7 and 9 years chronological age showed regular sequences of alpha waves, on the average, 54 percent of the time, while patients of the mongolian type of the same chronological age, but below MA 4, showed such sequences, on the average, only 17 percent of the time.

In a second report, Kreezer (116) extended his observations on a larger number of patients. He found a sudden break in his measures of alpha frequency, alpha amplitude, and percentage of regular alpha sequences occurring at MA 8 in the hereditary group of mentally deficient. Fifty subjects of the mongolian type, he adds in a footnote, presented evidence of a continuous change in electrocortical characteristics with increasing mental age, rather than a sudden shift at any one age level. Greater mean values of the three dimensions studied were found in the mongolian group; the differences were statistically significant, or nearly so.

Data secured from normal children ranging in age from three months to sixteen years have been presented by Lindsley (125). Alpha frequency showed a gradual increase with age, with the adult frequency attained at about ten years. Lindsley feels that the curve of increase in frequency is closely related to the curve of increase in brain weight. The measures of percentage of regular sequences apparently indicate a falling off with increase in age. This trend could not be predicted by any generalization from Kreezer's data; such a prediction would favor the opposite trend; that is, the percentage of regular sequences should increase with (mental) age.

The apparent disparity between the results of these two investigators may possibly rest on the essential difference of their populations. The low percentage of regular sequences in the mongolian and hereditary types of mental deficiency must have a lawful neural basis, which may, in turn, be related to the MA's of these particular subjects. Lindsley's data would indicate that the low percentage of regular sequences is not a function of MA, *per se*. It is interesting that Kreezer's subjects below MA 4 showed evidence of very slow (5 per second) waves. Such rhythms have been noted in deteriorated epileptics. It may be that, in the institutional cases, the low mental ages and the low percentages of regular sequences of alpha waves and the slow rhythms are all a function of some deteriorated condition. In the normal group, a mental age of four years or less would not be a function of such a condition. It may be that the problem of the disparity hinges upon whether or not MA 4 has the same significance in institutional feebleminded patients and in normal children.

In evaluating the whole problem of the relation between the EEG and intelligence, there seem to be at least three factors that have operated against the discovery of any correlation in normal subjects. First, the records have not been taken under conditions involving intelligent behavior. Such conditions might provide us with more meaningful data than have been uncovered to date. Second, techniques of analysis that are limited to certain characteristics of the alpha rhythm do not provide a complete analysis of the EEG. The alpha rhythm is only one of many rhythms or frequencies, and it should be considered only in relation to all of the others. The Grass wave-analyzer (83) might well be applied to this problem, for it analyzes the EEG as a function of amplitude *vs.* frequency over the whole frequency range (1 to 50 cycles). Third, there is the very real possibility that the EEG technique is incapable of providing an adequate and detailed picture of cortical function. This incapacity of the technique, however, has not yet been demonstrated.

XXI. The Relationship between Intelligence and the Internal Environment of the Brain

Common observation from the dawn of history indicates that sudden disturbance of brain circulation leads to unconsciousness or lack of mental effectiveness. Hill (94) lists many men like Descartes, who " buried his head in a sofa " in an effort to improve " thinking." These casual observations are said to suggest a correlation between excellence of mental function and a full supply of blood in optimal condition to the brain. Interference with the supply of blood to the brain almost instantaneously produces dizziness, as always noted when accidentally the main cerebral arteries are cut.

Many transient conditions influence the organism's ability to perform intelligent acts. By definition, one may say that conditions, such as sleep, that change the individual's intelligent reactions to his environment only temporarily are not factors related to ' intelligence ' as it is ordinarily considered, but a brief review of such factors is important for anyone who would understand intelligence, because they show, as it were in obverse, what the normal bodily conditions of intelligent behavior are. Also many apparently temporary conditions, such as acute oxygen deprivation, may permanently injure brain cells and leave the individual in a condition with respect to the performances of intelligent acts different from that in which he was before the apparently temporary ' unconscious episode.' Cannon (20) has pointed out that oxygen deprivation of 8 to 10 minutes may produce irreversible effects on cortical cells resulting in a permanent deficit in behavior. Other parts of the central nervous system are much more resistant to lack of oxygen than is the brain, and peripheral nerves are still less affected by such deprivation.

There seem, nevertheless, to be wide individual differences in the effect of compression of one carotid artery in different men, depending upon the inborn connections of that part of the cerebral blood system called the 'Circle of Willis.' Page (154) has recently pointed out in summarizing the literature on the chemistry of the brain that unconsciousness may result directly from deficient oxygen supply to the brain or to a spread of nervous inhibition that is itself caused by lack of oxygen. In this connection the work of Gibbs and Lennox (70) is also most important. Chronic anemia of the brain in childhood may lead to permanent imbecility.

Pernicious anemia is often accompanied by mental symptoms and the special mental phenomena of old age are known often to be related to changes in the circulation of blood in the brain (147).

Claude Bernard pointed out that " the constancy of the *milieu intérieur* is the condition of the free life." This statement summarizes the field studied by a number of modern physiologists. Cannon (19), Pike and Scott (163), Barcroft (4), Dill (33), Gellhorn (60, 61), McFarland (132), and many others have recently investigated certain aspects of the relation between the properties of the blood that make up the brain's 'internal environment' and the behavior of the individual. A résumé of this work as prepared by Barcroft is given in the following tabulation:

Environment	*Deficient*	*Excessive*
Temperature	Inertia	Delirium
Oxygen	Unconsciousness	
CO_2	Headache	Coma
Glucose	Nervousness Feeling of 'goneness' Hunger	
Water	Weakness (Asher)	Headache Nausea Dizziness Asthenia Incoördination

Environment	*Deficient*	*Excessive*
Sodium	Fever	Reflex irritability Weakness Paresis
Calcium	Nervous twitchings Convulsions	Apathy Drowsiness verging on coma General atonia

The influence of temperature upon mental life and man's intelligent behavior is given strikingly in the following excerpt from Barcroft (4) in which he describes certain experiments of his own:

> As I lay naked in the cold room at Woods Hole I had been shivering and my limbs had been flexed in a sort of effort to huddle up, and I had been very conscious of the cold. Then a moment came when I stretched out my legs, the sense of coldness passed away, and it was succeeded by a beautiful feeling of warmth; the word 'bask' most fitly describes my condition: I was basking in the cold. What had taken place, I suppose, was that my central nervous system had given up the fight, the vasoconstriction had passed from my skin, and the blood returning thither gave that sensation of warmth which one experiences when one goes out of a cold-storage room into the ordinary air. Perhaps I can express the change which took place in another way. Up to the point at which shivering ceased, nature fought the situation; my instinct was to be up and about, an effort of will was necessary to remain the subject of the experiment; after that point I gladly acquiesced, initiative had gone, and that is what I suppose Cannon to mean when he uses the word 'inertia.' Doubtless a second and more advanced stage would follow in which inertia would lapse into unconsciousness. For I suppose that, had the experiment not ended at that point, my temperature would have fallen rapidly and I was on the verge of the condition of travelers when they go to sleep in extreme cold never again to awake. (pp. 88–89)

The effect of increased temperature as reported by Sutton (193), involved, with an increase of 2 or 3 degrees, a loss in the ability to read a book or to learn by heart a short vocabulary of German words. Indeed, the most notable characteristic of the activity following increase in temperature was excessive excitement.

Especially because of its importance in determining conditions under which an airplane can be properly piloted at high altitudes, a great deal of study has been directed toward the temporary and permanent effect of the amount of oxygen in the air on human performance. In an important series of papers McFarland (129, 130, 131) reports experiments on mental changes resulting from change of altitude. In certain laboratory experiments the oxygen factor, as contrasted with other effects of altitude, has been differentially studied. The scientific work in this field is well summarized by McFarland (132) [1] and by Shock (182, 183). In connection with laboratory experiments McFarland reports:

> During the exposure to oxygen lack, varying in degree from 21 percent at sea level to 9 percent (corresponding roughly to 22,000 feet), the pilots were given a series of psychological tests, composed of simple and choice reaction times, perception of spacial relations, motor coördination, memory, judgment, and emotional control. When the oxygen deprivation was extended beyond simulated altitudes of 14,000 to 16,000 feet, the behavior of each pilot was profoundly altered. Not only was there a loss of judgment in relation to their own behavior and memory for recent events, but also there was a complete distortion of emotional control simulating many of the characteristics of certain mentally abnormal patients and of those suffering from excessive amounts of alcohol or narcotics. In the more acute experiments with 9 percent oxygen, the pilots would frequently lose the capacity for sane judgment and for self-criticism. Some of the pilots responded with great hilarity and uncontrollable laughter, while others became very angry and destructive. It was interesting that certain temperamental and emotional traits of each individual became greatly accentuated, suggesting that this procedure might be used to reveal certain basic aspects of behavior of psychiatric interest.

XXII. Other Factors, Especially Drugs, that Temporarily or Permanently Influence Human Intellectual Efficiency

Changes in the blood in respect to carbon dioxide and to blood sugar also affect the mental life of the individual, as of course do a whole series of anesthetics, hypnotics, narcotics, and other pharmacological substances. (See Refs. 49 and 84.) Shock (182, 183) has well summarized the effects of alcohol, aspirin, tobacco, caffeine, benzedrine, and phosphate upon mental efficiency. In most such cases, the problem of the acute action of the agent must be separated from the permanent condition that follows prolonged use; and in a consideration of the physiological basis of intelligence we are more concerned with the chronic than with the acute effects of such agents on adaptive

[1] The writer wishes to express his gratitude to Dr. R. A. McFarland for his kindness in making his bibliography on the psycho-physiological effects of reduced oxygen pressure available for the purposes of this chapter.

behavior. As already noted, however, there is growing evidence that some agents, such as oxygen lack, that lead to the dysfunction of brain cells, even for a relatively short time, may initiate an irreversible process in the protoplasm of the cell. In this connection, much will no doubt be learned in the next few years concerning the action of insulin and metrazol as used in shock therapies for certain types of mental illness. It has even been suggested speculatively that these drugs lead to the death of certain brain cells and thus physically disorganize the functional brain patterns that have been basic to the patient's mental illness.

Certain sedatives, such as luminal and dilantin, may change the patient's mental ability as shown by intelligence tests when they are given over a long period of time. Much research remains to be done in this field before it can be said with assurance how much of such deterioration results from the direct effect of the drug and how much from loss of functional exercise produced by the torpor that follows the use of the drug. The effect of cumulative toxic substances, like lead poisoning, has also been most extensively studied. In some cases of lead poisoning marked cerebral atrophy is present. In all such cases a loss of intelligence is reported.

A great deal of work has been done on the effect of diet and of the endocrines on mental development. This, which may be considered as falling under a general consideration of the effect of the internal environment on the brain, has been summarized by Dr. Nathan W. Shock. Section XXIII that follows is quoted by permission from an abstract of a paper read by Dr. Shock before the Society for Research in Child Development.

XXIII. The Effect of Benzedrine and of Thyroid and Pituitary Substances on Mental Life

Dr. Nathan W. Shock
Assistant Professor of Physiology
University of California
Berkeley, California

Alterations of the internal environment that result in increased intellectual performance are relatively rare. However, some recent studies of the effect of the drug, benzedrine, on mental performance are interesting.

Sargant and Blackburn (177) administered Cattell's test to twenty-five mental patients before and after the administration of twenty milligrams of benzedrine. A control group, which had received a blank tablet in place of the benzedrine, showed no change in average score, while the experimental group, which had received benzedrine, showed a rise of 3.9, or 8.7 percent, in the average score. The author interprets this increase in score in terms of the influence of non-intellectual factors, such as release from anxiety, mild depression, and the like. Similar increases in intellectual performance are reported by other authors in mentally retarded children of various ages. Molitch and Sullivan (150) found, for instance, that by increasing the dosage of benzedrine, significant increases in mental-test performance could be obtained in a total of 92 percent of a group of 96 boys between the ages of 10 and 17. A stimulating effect on motor activity of benzedrine has been reported in human beings and confirmed in actual measurements on animals. Although normal individuals report subjective feelings of increased mental efficiency and improved mental performance, McNamara and Miller (135) found that the number of problems of written multiplication of two three-digit numbers was not materially increased nor were the errors affected, even though the subjective feelings of stimulation were reported. On the other hand, the administration of benzedrine to psychiatric patients does not seem to affect the mood to any great extent. Hence the conclusion of Sargant and Blackburn, that the improvement in intelligence-test scores is due largely to nonintellectual factors, is open to question.

It has been shown experimentally with animals that quantitative limitations of the diet have little effect on learning ability, although vitamin-D deficiency has a definite deleterious influence on learning. Poull (166), however, in a study of 40 boys ranging in age from 3 years, 8 months, to 15 years, 7 months, found an average rise of 10 points in IQ when nutrition was improved between the first and second mental tests. Whether the improved nutrition can be considered as a cause of the increase in IQ from these experiments is questionable. The interesting possibility of inadequate dietary intake of inorganic phosphate contributing to fatigue, lassitude, and lowered intelligence of school children is raised by experiments carried on in Germany. Several clinical and laboratory experiments (165, 168, 190) have been carried out in which recresal or monosodium phosphate tends to reduce mental fatigue and increase intellectual output when administered to experimental subjects, particularly children. Since the effect of suggestion has in no way been controlled in most of these experiments, their true significance is uncertain.

The possibility of a relationship between endocrine function and mental performance has been an attractive one. Numerous experiments have shown the relationship between inadequate thyroid secretion and the incidence of retarded mental development. It has also been shown that thyroid feeding, if begun early enough and if continued with sufficient vigor, may result in considerable improvement in these cases. Since the basal metabolic rate is accepted as the clinical evaluation of the functional level and activity of the thyroid gland, it is reasonable to suppose that a correlation between intelligence and basal metabolic rate could be established. Hinton (95) reported sufficiently high correlations between Stanford-Binet mental-test scores and basal metabolism in a group of 30 orphanage children and 60 private-school children. He found a correlation of .736 between Stanford-Binet mental-test scores and basal metabolic rate, although the correlation was only .53 between Binet score and Arthur performance score on the same group of children. Other studies, including our own, have failed to corroborate these high correlations, and have found a correlation of the order of .20 to .30 between metabolic rate and intelligence scores. In summarizing this literature, one is forced to the conclusion that, although extreme deviations in activity of the thyroid gland may have an important influence on intellectual development, within the normal range of fluctuation, such a close association has not been established.

The relationship between pituitary function and intelligence or behavior is extremely problematic. Investigators, as, for instance, Mateer (141, 142, 143), report phenomenal improvement in reading ability resulting from pituitary administration. As is the case in many other clinical reports, educational, environmental, and glandular therapy were always administered simultaneously, so that a scientific evaluation of the results is impossible. Other individuals (139) report considerably less success with the administration of pituitary extract by mouth than from the administration of thyroid substance. In considering these clinical reports of the effect of pituitary therapy, it is important to remember that in no instance has a physiological effect of pituitary gland been observed in animal experiments where the pituitary substance has been administered by mouth. Whatever other disagreements may be found in the physiological literature on the pituitary gland, experimenters are practically unanimous in agreeing that all the active principles of the pituitary gland are destroyed in the alimentary tract, and hence are ineffective when administered by mouth (27, 43). Thus in evaluating the effect of pituitary medication on intelligence and behavior, we are faced with the dilemma of choosing either the rigorously controlled, but negative, results of animal experiments in the oral administration of pituitary extracts or the optimistic, but uncontrolled, reports of clinical success in the treatment of intellectual and behavior disorders.

XXIV. Conclusions and Suggested Educational Implications

Much else bearing upon the general topic of the relation between intelligent behavior and the functioning structures of the organ systems of the living body could be presented, but it is hoped that enough has been given to show something of the present status of research in this field. It is also hoped that the references given, which, though not complete, are reasonably comprehensive, will serve as an introduction to the literature of the field here discussed. I cannot help remarking that the lack of active scientific work in certain significant sections of this field is at once surprising and unfortunate. No civilization depends more upon the human brain and its ancillary organs than does ours. Yet for reasons that are not creditable to our social order too few first-class brains in each generation are induced to study the 'organ of human intelligence' itself. In certain important fields references a half century old still represent the last significant work.

The question may now be asked: What practical educational applications may be based upon this section of the Yearbook? In brief form, some of the possible applications are now presented. Not all of the statements appearing here below in dogmatic form have been fully defended in the foregoing presentation; some represent merely current opinions on the subjects treated, as I interpret such opinions. Every positive statement in so complex a field as the one here dealt with is open to at least some qualifications.

1. Intelligent behavior and the mental processes with which education is concerned have never been demonstrated to take place in the absence of active bodily structures. In the mammals and man such behavior is most directly related to the neocortex of the cerebral hemispheres of the brain. These portions of the brain most strictly involved in the cortical language mechanisms and the frontal lobes of the cerebrum are possibly of unique importance in at least certain types of characteristic intelligent behavior.

2. Direct anatomical, histological, physiological, or biochemical study of the brain or the nervous system does not serve — save in certain extreme cases of mental deficiency; for example, the brains of microcephalic aments — to make possible, in the absence of data derived from behavior, any judgment concerning the comparative intellectual status of an individual. The anatomical study of the brains of men and women of distinguished intellectual attainments has not yet shown why these individuals were better in intelligence than their merely average fellows.

3. There are excellent reasons to believe that the general anatomical characteristics of the brain of each human being are determined by processes that must be attributed to what is loosely called 'heredity' rather than to any specific set of 'environmental' circumstances, other than those essential to normal growth and health.

4. Behavior is a function of the make-up of the total organism,

including the brain; hence changes in intelligent behavior during infancy, childhood, the middle segment of life, or senescence are attributable to the working out of differential factors that are, in the sense indicated in the last paragraph, to be assigned to 'maturation' or, in its widest connotation, to inheritance. There seems to be little doubt that, given the same or comparable environments, brain characteristics, and hence characteristics of intelligent behavior, will change during growth. These growth changes come as a result of nerve-cell alterations, circulatory modifications, and the like, that are determined by factors intrinsic to the individual organism and that are, in the last analysis, dependent upon the specific inherited chromosomal background of that organism.

5. Environment, when considered as the sum of specific stimuli, releases much or — according to definition — all of the behavior of the living individual. Educational procedures are most important parts of specific environments. Thus, the individual 'educational environment' provides each human being with the behavior patterns that make possible the particular 'symbols,' 'concepts,' and 'ideas' he must employ in intelligent acts. All brains, for example, capable of developing symbolic processes of the sort necessary in involved mathematical operations do not come to be able to mediate the behavior necessary to solve problems dependent upon the use of higher mathematics — this because the environment may not provide the individual in question with that particular segment of social heritage called mathematics. Hence, the specific educational regimen provided for each child is of the greatest importance in relation to the intelligence he will demonstrate in any specific social order as measured by tests devised for use in a given society; that is, there is no physiological or psychological reason why educational procedures may not so modify each individual organism that he will be more effective in solving certain types of problems than he would be without the specific education in question. This statement does not necessarily mean that specific practice has changed the basic characteristics of the masses of organized nerve cells involved, or indeed of specific cells, but only that the given cell masses have been so modified in the process of learning that linguistic and other symbolic processes of the type possible in the brain protoplasm of the individual under consideration have been developed to a certain level. This means that, after specific education, behavior at that level of intelligence thus becomes possible when it would not be possible without such education. This statement is not

intended to imply any type of 'transfer of training' not fully substantiated by modern experiment.

6. Society in general, and educators in particular, can thus be encouraged to provide the richest possible environment in order to allow each brain in each generation to acquire to the maximal extent possible in that particular brain those 'symbols,' 'concepts,' 'ideas,' and 'habits of work' that will be useful to the individual in the problems he will meet in a life lived in the social system into whose membership he is growing.

7. Society in general, and educators in particular, would do well to promote a really scientific understanding of eugenics or proper parental selection for the children of each generation. If this is done, society may in time learn to avoid some of those breeding procedures that most often perpetuate unsatisfactory brains and inevitably lead to behavior that is defective in those acts called intelligent in any given society. In this connection, it must be recognized, however, that much popular eugenics is not sound when checked against the modern quantitative knowledge amassed by the special biological science of genetics.

8. Since also many environmental factors, such as nourishment, especially including the appropriate provision or exclusion of the vitamins, oxygen, toxic substances, disease products in the blood stream, and the like, are known to influence the development of the brain, whatever its hereditary 'potentiality' may be, such factors should be controlled to the fullest degree possible by a society that wishes to maintain the maximal intelligence in the behavior of its population. Thus, every effort may well be made by society and by educators to provide the most adequate human nourishment in its completest sense in an environment as free as possible from toxic agents in order to influence the optimal growth of each brain and thus to allow the optimal education for each human brain within the boundary set by the inherited anatomical limits of each brain.

9. What has just been said concerning the interdependent relationship of heredity and environment in the determination of the bodily basis of intelligent behavior throws important light on the sometimes divergent conclusions of scientific studies on the change of the 'intelligence quotient' during individual growth as determined by intelligence tests. Such tests depend at least in part upon the measurement of behavior involving the 'symbols,' 'concepts,' and 'ideas' of specific social order. If we assume that studies using such tests are properly controlled and that they are evaluated by appropriate statistical pro-

cedures, there is still room for variation in the magnitude of the change of IQ produced by superficially similar environments. Such differences may be found to be dependent upon real, but often subtle, environmental peculiarities, such as those summarized in the word 'motivation,' as well as by the more gross factors of nutrition and the formal nature of the educational curriculum offered to the groups being compared.

References

(1) Alford, L. B. "Defects of intelligence from focal lesions within the central part of the left cerebral hemisphere." *Amer. Jour. Psychiat.*, 94: 1937, 615–638.

(2) Angulo y González, A. W. "Is myelinogeny an absolute index of behavioral capability?" *Jour. Comp. Neur.*, 48: 1929, 459–464.

(3) Anthony, R. *Anatomie comparée du cerveau.* (Doin: Paris, 1928) 359 pp.

(4) Barcroft, J. *The Brain and its Environment.* (Yale University Press: New Haven, 1938) 117 pp.

(5) Bard, P., and Rioch, D. M. "A study of four cats deprived of neocortex and additional portions of the forebrain." *Bull. Johns Hopkins Hospital,* 60: 1937, 73–147.

(6) Bastian, H. C. *The Brain as an Organ of Mind.* (D. Appleton-Century: New York, 1880) 708 pp.

(7) Bell, Sir C. "The anatomy of the brain." [In] Spurzheim, J. C. *The Anatomy of the Brain.* (Marsh, Capen and Lyon: Boston, 1836) pp. 217–244.

(8) Berger, H. "Über das Elektrenkephalogramm des Menschen: I." *Arch. f. Psychiat.*, 87: 1929, 527–570.

(9) Berger, H. "Über das Elektrenkephalogramm des Menschen: IV." *Arch. f. Psychiat.*, 97: 1932, 6–26.

(10) Berger, H. "Über das Elektrenkephalogramm des Menschen: V." *Arch. f. Psychiat.*, 98: 1933, 232–251.

(11) Bigelow, H. J. "Dr. Harlow's case of recovery from the passage of an iron bar through the head." *Amer. Jour. Med Sciences,* New Series 20: 1850, 13–22.

(12) Bischoff, T. L. W. *Das Hirngewicht des Menschen.* (Neusser: Bonn, 1880) 171 pp.

(13) Bolton, J. S. *The Cortical Localization of Cerebral Function.* (Oliver and Boyd: Edinburgh, 1933) 23 pp. (The Henderson Trust Lectures No. XII)

(14) Bonin, G. v. "Brain-weight and body-weight of mammals." *Jour. Gen. Psychol.*, 16: 1937, 379–389.

(15) Brickner, R. M. "An interpretation of frontal lobe function based upon the study of a case of partial bilateral frontal lobectomy." [In] *Localization of Function in the Cerebral Cortex.* Association for Research in Nervous and Mental Disease. (Williams and Wilkins: Baltimore, 1934) (Series of Research Publications, Vol. XIII, pp. 259–351)

(16) BRICKNER, R. M. *The Intellectual Functions of the Frontal Lobes.* (Macmillan: New York, 1936) 354 pp.

(17) BURR, H. S. "Determinants of organization in the cerebral hemispheres." [In] *Localization of Function in the Cerebral Cortex.* Assoc. for Research in Nervous and Mental Disease. (Williams and Wilkins: Baltimore, 1934) (Series of Research Publications, Vol. XIII)

(18) CAMERON, N. "Cerebral destruction in its relation to maze learning." *Psychol. Monog.,* 39: 1928–1929, 1–68.

(19) CANNON, W. B. *Bodily Changes in Pain, Hunger, Fear, and Rage.* (D. Appleton-Century: New York, 2d ed.) 404 pp.

(20) CANNON, W. B. "Organization for physiological homeostasis." *Physiological Reviews,* 9: 1929, 399–431.

(21) CARLYLE, F., JACOBSEN, M. M., and YOSHIOKA, J. G. [For corrected reference, *see* Jacobsen, C. F., *et al,* 102a.]

(22) CARMICHAEL, L. "Origin and prenatal growth of behavior." [In] Murchison, C., editor. *Handbook of Child Psychology,* 2nd rev. ed. (Clark University Press: Worcester, 1933) pp. 31–159.

(23) CARMICHAEL, L. "The psychology of genius." *Phi Kappa Phi Jour.,* 1934, 149–164.

(24) CARMICHAEL, L. "Fetal behavior and developmental psychology." *Onzième Congrès Internationale de Psychologie,* 1938, 108–117.

(25) CARMICHAEL, L. "Chapter VII. The response mechanism." [In] Boring, E. G., Langfeld, H. S., and Weld, H. P., editors. *Introduction to Psychology.* (John Wiley: New York, 1939) pp. 223–249.

(26) COGHILL, G. E. "The neuro-embryologic study of behavior: Principles, perspective and aim." *Science,* 78: 1933, 131–138.

(27) COLLIP, J. B. "Diabetogenic, thyrotropic, adrenotropic and parathyrotropic factors of the pituitary: A symposium prepared under the auspices of the Council on Pharmacy and Chemistry of the American Medical Association." [In] *Glandular Physiology and Therapy.* (American Medical Association: Chicago, 1935) Chap. VII, pp. 85–117.

(28) DANDY, W. E. "Changes in our conceptions of localization of certain functions in the brain." Abstract. *Amer. Jour. Physiol.,* 1930, 93, 643.

(29) DAVIS, H. "The electroencephalogram." *Tabulae Biologicae,* 16: 1938, 116–131.

(30) DAVIS, H., and DAVIS, P. A. "Action potentials of the brain in normal persons and in normal states of cerebral activity." *Arch. Neurol. and Psychiat.,* 36: 1936, 1214–1224.

(31) DEWEY, E. *Behavior Development in Infants.* (Columbia University Press: New York, 1935) 321 pp.

(32) DILL, D. B. *Life, Heat, and Altitude: Physiological Effects of Hot Climates and Great Heights.* (Harvard University Press: Cambridge, 1938) 211 pp.

(33) DOLL, E. A. "Birth lesion as a category of mental deficiency." *Amer. Jour. Orthopsychiat.,* 3: 1933, 1–13.

(34) DOLL, E. A. "Psychological significance of cerebral birth lesions." *Amer. Jour. Psychol.,* 45: 1933, 444–452.

(35) DOLL, E. A., PHELPS, W. M., and MELCHER, R. T. *Mental Deficiency Due to Birth Injuries.* (Macmillan: New York, 1932) 289 pp.

(36) DONALDSON, H. H. *The Growth of the Brain: A Study of the Nervous System in Relation to Education.* (Walter Scott: London, 1895) 374 pp.

(37) DUBOIS, E. "Über die Abhängigkeit des Hirngewichtes von der Körpergrösse beim Menschen." *Arch. f. Anthropol.*, 25: 1898, 1–28.

(38) DUGDALE, R. L. *The Jukes: A Study in Crime, Pauperism, Disease, and Heredity.* (Putnam: New York, 1910, 4th ed.) 120 pp.

(39) DUSSER DE BARENNE, J. G. "Some aspects of the problem of 'corticalization' of function and of functional localization in the cerebral cortex." [In] *Localization of Function in the Cerebral Cortex.* (Williams and Wilkins: Baltimore, 1934) (Series of Research Publications, Vol. XIII, pp. 85–106)

(40) EHRENFEST, H. *Birth Injuries of the Child.* (D. Appleton-Century: New York, 1931, 2d ed.) 317 pp.

(41) EPSTEIN, S. H., and SOLOMON, H. C. "The effect of treatment on the mental level of patients with general paresis." *Amer. Jour. Psychiat.*, 95: 1939, 1181–1192.

(42) ERICKSEN, S. C. "The relative effect of a cerebral lesion upon learning, retention, and transfer." *Jour. Comp. Psychol.*, 27: 1939, 373–391.

(43) EVANS, H. M. "The growth hormone of the anterior pituitary." [In] *Glandular Physiology Therapy: A Symposium Prepared under the Auspices of the Council on Pharmacy and Chemistry of the American Medical Association.* (Amer. Med. Assoc.: Chicago, 1935, Chap. III, pp. 45–59)

(44) FERRIER, D. *The Functions of the Brain.* (Smith, Elder: London, 1886. 2nd ed.) 498 pp.

(45) FISCHEL, M. K. *The Spastic Child: A Record of Successfully Achieved Muscle Control in Little's Disease.* (C. V. Mosby: St. Louis, 1934) 97 pp.

(46) FISHER, R. A. *Statistical Methods for Research Workers.* (Oliver and Boyd: Edinburgh, 1930. 3rd. ed.) 283 pp.

(47) FLECHSIG, P. *Die Leitungsbahnen im Gehirn und Rückenmark des Menschen auf Grund entwicklungsgeschichtlicher Untersuchungen.* (Wilhelm Engelmann: Leipzig, 1879) 382 pp.

(48) FLECHSIG, P. *Gehirn und Seele.* (Veit: Leipzig, 1896) 112 pp.

(49) FORBES, A., and MORISON, B. R. "Cortical response to sensory stimulation under deep barbiturate narcosis." *Jour. Neurophysiol.*, 2: 1939, 112–128.

(50) FORD, F. R., CROTHERS, B., and PUTNAM, M. C. *Birth Lesions of the Central Nervous System.* (Williams and Wilkins: Baltimore, 1927) 164 pp.

(51) FOSTER, SIR M. *Lectures on the History of Physiology during the Sixteenth, Seventeenth, and Eighteenth Centuries.* (University Press: Cambridge, 1901) 310 pp.

(52) FRANZ, S. I. "New phrenology." *Science*, 35: 1912, 321–328.

(53) FRANZ, S. I. "Cerebral-mental relations." *Psychol. Rev.*, 28: 1921, 81–95.

(54) FRANZ, S. I. *The Evolution of an Idea: How the Brain Works.* (University of California: Los Angeles, 1929) 35 pp.

(55) FRANZ, S. I. "The neurology of learning." [In] Moss, F. A., editor. *Comparative Psychology.* (Prentice-Hall: New York, 1934) pp. 219–245.

(56) FRAZIER, C. H. "Tumor involving the frontal lobe alone: Symptomatic survey of 105 verified cases." *Arch. Neurol. and Psychiat.*, 35: 1936, 525–671.

(57) FROLOV, Y. P. *Pavlov and His School.* (K. Paul, Trench, Trubner, and Co.: London, 1937) 291 pp.

(58) FULTON, J. F. *Physiology of the Nervous System.* (Oxford University Press: London, 1938) 675 pp.

(59) GALL, F. J. *On the Functions of the Brain and Each of its Parts.* Works of Franz Joseph Gall, trans. from the French by Winslow Lewis. (Marsh, Capen and Lyon: Boston, 1835) Vols. 1–6.

(60) GELLHORN, ERNST. "The effect of O_2-lack variations in the CO_2-content of the inspired air, and hyperpnea on visual intensity discrimination." *Amer. Jour. Physiol.*, 115: 1936, 679–684.

(61) GELLHORN, E. "The influence of carbon dioxide in combating the effect of oxygen deficiency on psychic processes with remarks of the fundamental relationship between psychic and physiologic reactions." *Amer. Jour. Psychiat.*, 93: 1937, 1413–1434.

(62) GERMAN, W. J., and FOX, J. C. "Observations following unilateral lobectomies." [In] *Localization of Function in the Cerebral Cortex.* (Williams and Wilkins: Baltimore, 1934) (Series of research publications, Vol. XIII, pp. 378–434)

(63) GESELL, A. "Motor disability and mental growth: The psychological effects of a cerebral birth palsy." *Psychol. Record,* 1: 1937, 87–94.

(64) GESELL, A., THOMPSON, H., and AMATRUDA, C. S. *Infant Behavior: Its Genesis and Growth.* (McGraw-Hill: New York, 1934) 333 pp.

(65) GESELL, A., and THOMPSON, H. "Learning and growth in identical infant twins: An experimental study by the method of co-twin control." *Genet. Psychol. Monog,* 6: 1929, 1–124.

(66) GESELL, A., AMATRUDA, C. S., CASTNER, B. M., and THOMPSON, H. *Biographies of Child Development; the Mental Growth and Careers of Eighty-Four Infants and Children.* (P. B. Hoeber, Inc.: New York, 1939) 328 pp.

(67) GESELL, A., and ZIMMERMAN, H. M. "Correlations of behavior and neuropathology in a case of cerebral palsy from birth injury." *Amer. Jour. Psychiat.*, 94: 1937, 505–535.

(68) GHISELLI, E. E., and BROWN, C. W. "Subcortical mechanisms in learning, VII." *Jour. Comp. Psychol.*, 26: 1938, 301–309.

(69) GIBBS, F. A. "Frequency with which tumors in various parts of the brain produce certain symptoms." *Arch. Neurol. and Psychiat.*, 28: 1932, 969–989.

(70) GIBBS, F. A., GIBBS, E. L., and LENNOX, W. G. "Changes in human cerebral blood flow consequent on alterations in blood gases." *Amer. Jour. Physiol.*, 111: 1935, 557–563.

(71) GIRDEN, E., METTLER, F. A., FINCH, G., and CULLER, E. "Conditioned responses in a decorticate dog to acoustic, thermal, and tactile stimulation." *Jour. Comp. Psychol.*, 21: 1936, 367–385.

(72) GODDARD, H. H. "The educational treatment of the feeble-minded." [In] WHITE, W. A., and JELLIFFE, S. E. *The Modern Treatment of Nervous and Mental Diseases.* (Lea and Febiger: Philadelphia, 1913) pp. 143–240.

(73) GODDARD, H. H. *Feeble-Mindedness: Its Causes and Consequences.* (Macmillan: New York, 1914) 599 pp.

(74) GODDARD, H. H. "The menace of mental deficiency from the standpoint of heredity." *Boston Med. and Surg. Jour.*, 175: 1916, 269–271.

(75) GODDARD, H. H. *The Kallikak Family: A Study in the Heredity of Feeble-Mindedness.* (Macmillan: New York, 1916) 121 pp.

(76) GOLDSTEIN, K. "Die Lokalisation in der Grosshirnrinde nach den Erfahrungen am kranken Menschen." [In] *Handb. der Normalen und Pathologischen Physiologie.* (Springer: Berlin, 1927) Vol. X, pp. 600–842.

(77) GOLDSTEIN, K. "The modifications of behavior consequent to cerebral lesions." *Psychiat. Quart.*, 10: 1936, 586–610.

(78) GOLDSTEIN, K. "The problem of the meaning of words, based upon observation of aphasic patients." *Jour. Psychol.*, 2: 1936, 301–316.

(79) GOLDSTEIN, K. "The significance of the frontal lobes for mental performances." *Jour. Neurol. Psychopath.*, 17: 1936, 27–40.

(80) GOLDSTEIN, K. "Psychologische Analysen hirnpathologischer Faelle." *Ztsch. f. d. ges. Neur. u. Psychiat.*, 41: 1918. Partially translated in Ellis, Willis Davis, *A Source Book of Gestalt Psychology.* (K. Paul, Trench, Trubner, and Co.: London, 1938)

(81) GOLDSTEIN, K., and BOLLES, M. "A study of the impairment of 'abstract behavior' in schizophrenic patients." *Psychiat. Quart.*, 12: 1938, 42–65.

(82) GOLDSTEIN, K. *The Organism: A Holistic Approach to Biology Derived from Pathological Data in Man.* (American Book Co.: New York, 1939) 533 pp.

(83) GRASS, A. M., and GIBBS, F. A. "A Fourier transformer of the electro-encephalogram." *Jour. Neurophysiol.*, 1: 1938, 521–526.

(84) GWATHMEY, J. T. *Anesthesia.* (Macmillan: New York, 1924, 2d ed.) 799 pp.

(85) HARLOW, J. M. "Passage of an iron rod through the head." *Boston Med. and Surg. Jour.*, 39: 1848, 389–393.

(86) HEAD, H. *Studies in Neurology.* (Frowde, Hodder, and Stoughton: London, 1920) 2 vols., 329 pp., 333–862 pp.

(87) HEAD, H. *Aphasia and Kindred Disorders of Speech.* (Macmillan: New York, 1926) Vol. 1, 549 pp.; Vol. 2, 430 pp.

(88) HENRY, G. W. "Mental phenomena observed in cases of brain tumor." *Amer. Jour. Psychiat.*, 12: 1932, 415–473.

(89) HENSCHEN, S. E. "Behandlung der Erkrankungen des Gehirns." [In] *Handbuch der Therapie innerer Krankheiten.* (Jena: Fischer, 1898) Vol. VI, 179 pp.

(90) HERON, W. T. "The inheritance of maze-learning ability in rats." *Jour. Comp. Psychol.*, 19: 1935, 77–89.

(91) HERRICK, C. J. *An Introduction to Neurology.* (Saunders: Philadelphia, 1921, 2d ed.) 394 pp.

(92) HERRICK, C. J. *Neurological Foundations of Animal Behavior.* (Holt: New York, 1924) 334 pp.

(93) HERRICK, C. J. *Brains of Rats and Men: A Survey of the Origin and Biological Significance of the Cerebral Cortex.* (University of Chicago Press: Chicago, 1926) 382 pp.

(94) HILL, L. *The Physiology and Pathology of the Cerebral Circulation: An Experimental Research.* (London, 1896) 208 pp.

(95) HINTON, R. T., JR. "A further study of the rôle of the basal metabolic rate in the intelligence of children." *Jour. Educ. Psychol.*, 30: 1939, 309–314.

(96) Hollander, B. *The Revival of Phrenology.* (G. Putnam: London, 1901) 512 pp.

(97) Holmes, G. "Discussion on the mental symptoms associated with cerebral tumors." *Proc. Roy. Soc. Med.*, 24: 1931, 997–1000.

(98) Hooker, D. "Early fetal activity in mammals." *Yale Jour. Biol. and Med.*, 8: 1936, 579–602.

(99) Hooker, D. "The origin of the grasping movement in man." *Proc. Amer. Phil. Soc.*, 79: 1938, 597–606.

(100) Hoopes, G. G. *Out of the Running.* (Charles C. Thomas: Springfield, 1939) 158 pp.

(101) Jacobsen, C. F. "A study of the cerebral function in learning: The frontal lobes." *Jour. Comp. Neurol.*, 52: 1931, 271–340.

(102) Jacobsen, C. F. "Studies of cerebral function in primates." *Comp. Psychol. Monog.*, 13: 1936, 1–68.

(102a) Jacobsen, C. F., Jacobsen, M. M., and Yoshioka, J. G. "Development of an infant chimpanzee during her first year." *Psychol. Monog.*, 9: 1932, 1–94.

(103) Jacobsen, C. F., Wolfe, J. B., and Jackson, T. A. "An experimental analysis of the functions of the frontal association areas in primates." *Jour. Nervous and Mental Disease*, 82: 1935, 1–14.

(104) Jasper, H. H. "Electrical signs of cortical activity." *Psychol. Bull.*, 34: 1937, 411–481.

(105) Jasper, H. H., Bridgman, C. S., and Carmichael, L. "An ontogenetic study of cerebral electrical potentials in the guinea pig." *Jour. Exper. Psychol.*, 21: 1937, 63–71.

(106) Jefferson, G. "Removal of right or left frontal lobes in man." *Brit. Med. Jour.*, 2: 1937, 199–206.

(107) Kappers, C. V. A., Huber, G. C., and Crosby, E. C. *The Comparative Anatomy of the Nervous System of Vertebrates, including Man.* (Macmillan: New York, 1936) 2 vols., 1845 pp.

(108) Kasatkin, N. I., and Levikova, A. M. "The formation of visual conditioned reflexes and their differentiation in infants." *Jour Gen. Psychol.*, 12: 1935, 416–435.

(109) Kasatkin, N. I., and Levikova, A. M. "On the development of early conditioned reflexes and differentiations of auditory stimuli in infants." *Jour. Exper. Psychol.*, 18: 1935, 1–19.

(110) Kellogg, W. N., and Kellogg, L. A. *The Ape and the Child: A Study of Environmental Influence upon Early Behavior.* (Whittlesey, McGraw-Hill: New York, 1933) 314 pp.

(111) Kennedy, F. "The symptomatology of frontal and temporosphenoidal tumors." *Jour. Amer. Med. Assoc.*, 98: 1932, 864–866.

(112) Keschner, M., Bender, M. B., and Strauss, I. "Mental symptoms in cases of tumor of the temporal lobe." *Arch. Neurol. and Psychiat.*, 35: 1936, 572–596.

(113) Koffka, K. (Trans. by Ogden, R. M.) *The Growth of the Mind: An Introduction to Child Psychology.* (Harcourt, Brace: New York, 1925) 382 pp.

(114) Krechevsky, I. "Brain mechanism and 'hypotheses.'" *Jour. Comp. Psychol.*, 19: 1935, 425–468.

(115) KREEZER, G. " Electric potentials of the brain in certain types of mental deficiency." *Arch. Neurol. and Psychiat.*, 36: 1936, 1206–1213.

(116) KREEZER, G. " Electrical phenomena of the brain among the feeble-minded." *Proc. Amer. Assoc. Mental Deficiency*, 42: 1937, 130–141.

(117) KREEZER, G. " The dependence of the electroencephalogram upon intelligence level." *Psychol. Bull.*, 34: 1937, 769.

(118) KREEZER, G. " The electroencephalogram and its use in psychology." *Amer. Jour. Psychol.*, 51: 1938, 737–759.

(118a) KREEZER, G. " Intelligence level and occipital alpha rhythm in the Mongolian type of mental deficiency." *Amer. Jour. Psychol.*, 52: 1939, 503–532.

(119) LAPICQUE, L. " Le Poids du Cerveau et l'Intelligence." [In] DUMAS, G., *Nouveau Traité de Psychologie*. (Felix Alcan: Paris, 1930) Vol. I, 204–218.

(120) LASHLEY, K. S. *Brain Mechanisms and Intelligence: A Quantitative Study of Injuries to the Brain.* (University of Chicago Press: Chicago, 1929) 186 pp.

(121) LASHLEY, K. S. " Experimental analysis of instinctive behavior." *Psychol. Rev.*, 45: 1938, 445–471.

(122) LASHLEY, K. S. " Factors limiting recovery after central nervous lesions." *Jour. Nervous and Mental Disease*, 88: 1938, 733–755.

(123) LASHLEY, K. S., and WILEY, L. E. " Studies of cerebral function in learning. IX. Mass action in relation to the number of elements in the problem to be learned." *Jour. Comp. Neurol.*, 57: 1933, 3–55.

(124) LINDSLEY, D. B. " Brain potentials in children and adults." *Science*, 84: 1936, 354.

(125) LINDSLEY, D. B. " Electrical potentials of the brain in children and adults." *Jour. Gen. Psychol.*, 19: 1938, 285–306.

(125a) LINDSLEY, D. B. " A longitudinal study of the occipital alpha rhythm in normal children: Frequency and amplitude standards." *Jour. Genet. Psychol.*, 55: 1939, 197–213.

(126) LOOMIS, A. L., HARVEY, E. N., and HOBART, G. " Electrical potentials of the human brain." *Jour. Exper. Psychol.*, 19: 1936, 249–279.

(127) LORD, E. E. *Children Handicapped by Cerebral Palsy: Psychological Factors in Management.* (The Commonwealth Fund: New York, 1937) 105 pp.

(128) McDOUGALL, W. *Body and Mind: A History and a Defense of Animism.* (Methuen: London, 6th ed. 1923) 384 pp.

(129) McFARLAND, R. A. " The psychological effects of oxygen deprivation (anoxemia) on human behavior." *Arch. Psychol.*, 22: 1932, 5–135.

(130) McFARLAND, R. A. " Psycho-physiological studies at high altitudes in the Andes. II. Sensory and motor responses during acclimatization. *Jour. Comp. Psychol.*, 23: 1937, 227–258.

(131) McFARLAND, R. A. " Psycho-physiological studies at high altitudes in the Andes. III. Mental and psycho-somatic responses during gradual adaptation." *Jour. Comp. Psychol.*, 24: 1937, 147–188.

(132) McFARLAND, R. A. " The psycho-physiological effects of reduced oxygen pressure." [To appear in] *Proceedings of the Association for Research in Nervous and Mental Diseases.*

(133) McGRAW, MYRTLE B. *Growth: A Study of Johnny and Jimmy.* (D. Appleton-Century: New York, 1935) 316 pp.

(134) McIntire, J. T. "The incidence of feeble-mindedness in the cerebral palsied." *Proc. Amer. Assoc. Mental Deficiency,* 43: 1938, 44–50.

(135) McNamara, W. J., and Miller, R. E. "Effect of benzedrine sulphate on mental work." *Psychol. Record,* 1: 1937, 78–84.

(136) Maier, N. R. F. "The effect of cerebral destruction on reasoning and learning in rats." *Jour. Comp. Neurol.,* 54: 1932, 45–75.

(137) Maier, N. R. F. "Cortical destruction of the posterior part of the brain and its effect on reasoning in rats." *Jour. Comp. Neurol.,* 56: 1932, 179–214.

(138) Maier, N. R. F. "Qualitative differences in the problem-solving of normal and partly decorticated rats." *Mich. Acad. Science, Arts and Letters,* 24: 1939, 131–146.

(139) Marinus, C. J., and Kimball, O. P. "Endocrine dysfunctions in retarded children and their response to treatment." *Endocrinology,* 14: 1930, 309–318.

(140) Marquis, D. P. "Can conditioned responses be established in the newborn infant?" *Ped. Sem. and Jour. Genet. Psychol.,* 39: 1931, 479–492.

(141) Mateer, F. "Pituitary acceleration of IQ's." *Psychol. Bull.,* 31: 1934, 686.

(142) Mateer, F. "A first study of pituitary dysfunction in cases of reading difficulty." *Psychol. Bull.,* 32: 1935, 736.

(143) Mateer, F. *Glands and Efficient Behavior.* (D. Appleton-Century: New York, 1935) 243 pp.

(144) Melton, A. W. "A note on the relation between brain injury and retentiveness." *Psychol. Bull.,* 28: 1931, 175–180.

(145) Mettler, F. A. "Cerebral function and cortical localization." *Jour. Gen. Psychol.,* 13: 1935, 367–401.

(146) Mettler, F. A. "The present status of the frontal lobe." *Jour. Gen. Psychol.,* 17: 1937, 203–233.

(147) Miles, W. R. "Psychological aspects of ageing." [In] *Problems of Ageing.* Cowdry, E. V., editor. (Williams and Wilkins: Baltimore, 1939) pp. 535–571.

(148) Minkowski, M. "Zur Entwicklungsgeschichte, Lokalisation, und Klinik des Fussohlenreflexes." *Schweizer Arch. f. Neurol. und Psychiat.,* 13: 1923, 475–514.

(149) Minkowski, M. "Neurobiologische Studien am menschlichen Foetus." *Abderhaldens Hdbh. d. biol. Arbeitsmeth.,* 1928, 511–618.

(150) Molitch, M., and Sullivan, J. P. "The effect of benzedrine sulfate on children taking the new Stanford Achievement Test." *Amer. Jour. Orthopsychiat.,* 7: 1937, 519–522.

(151) Monakow, C. V. *Die Lokalisation im Grosshirn und der Abbau der Funktion durch kortikale Herde.* (Bergmann: Wiesbaden, 1914) 1033 pp.

(152) Munn, N. L. *Psychological Development: An Introduction to Genetic Psychology.* (Houghton Mifflin: Boston, 1938) 582 pp.

(153) Newman, H. H., Freeman, F. N., and Holzinger, K. J. *Twins: A Study of Heredity and Environment.* (University of Chicago Press: Chicago, 1937) 369 pp.

(154) Page, I. H. *Chemistry of the Brain.* (Thomas: Springfield, 1937) 440 pp.

(155) Paterson, D. G. *Physique and Intellect.* (D. Appleton-Century: New York, 1930) 304 pp.

(156) PEIPER, A. *Die Hirntätigkeit des Säuglings.* (Springer: Berlin, 1928) 102 pp.
(157) PENFIELD, W., and EVANS J. "The frontal lobe in man: A clinical study of maximum removals." *Brain,* 58: 1935, 115–133.
(158) PESSIN, J. "Mental symptoms in brain tumor." *Jour. Kansas Med. Soc.,* 38: 1937, 513–524.
(159) PHELPS, W. M. "The care and treatment of cerebral palsies." *Jour. Amer. Med. Assoc.,* 111: 1938, 1–6.
(160) PIÉRON, H. *Thought and the Brain.* Trans. by Ogden, C. K. (Harcourt Brace: New York, 1927) 262 pp.
(161) PIKE, F. H. "Remarks on von Monakow's 'Die Lokalisation im Grosshirn.'" *Jour. Comp. Neurol.,* 29: 1918, 485–508.
(162) PIKE, F. H., and CHAPPELL, M. N. "On the recovery following lesions in the cerebral cortex." *Science,* 71: 1930, 76.
(163) PIKE, F. H., and SCOTT, E. L. "The significance of certain internal conditions of the organism in organic evolution." *Amer. Naturalist,* 49: 1915, 321–359.
(164) PIOTROWSKI, Z. "The Rorschach inkblot method in organic disturbances of the central nervous system." *Jour. Nerv. and Ment. Disease,* 86: 1937, 525–537.
(165) POPPELREUTER, W. "Steigerung der geistigen Arbeitsfähigkeit durch Recresalzufuhr." *München. med. Wochensch.,* 76: 1929, 912–915.
(166) POULL, L. E. "The effect of improvement in nutrition on the mental capacity of young children." *Child Develop.,* 9: 1938, 123–126.
(167) PRATT, K. C. "The neonate." [In] Murchison, C., editor. *Handbook of Child Psychology.* (Clark University Press: Worcester, 1933) 2d ed., pp. 163–208.
(168) RAHM, K. "Über die Wirkung des Recresals auf die körperliche und geistige Leistungsfähigkeit." *Arch. f. die ges. Psychol.,* 86: 1932, 459–522.
(169) RAHM, W. E., and WILLIAMS, A. C. "Aspects of the electroencephalogram in epilepsy and feeble-mindedness." *Psychiat. Quart.,* 12: 1938, 230–235.
(170) RANSON, S. W. *The Anatomy of the Nervous System from the Standpoint of Development and Function.* (Saunders: Philadelphia, 1935) 5th ed., 501 pp.
(171) RAY, W. S. "A preliminary report on a study of fetal conditioning." *Child Develop.,* 3: 1932, 175–177.
(172) RETZIUS, G. "Das Gehirn des Mathematikers Sonja Kovalevski." *Biologische Untersuchungen,* 9: 1900, 1–16.
(173) ROGERS, G. G., and THOMAS, L. C. *New Pathways for Children with Cerebral Palsy.* (Macmillan: New York, 1935) 167 pp.
(174) ROWE, S. N. "Mental changes following the removal of the right cerebral hemisphere for brain tumor." *Amer. Jour. Psychiat.,* 94: 1937, 605–614.
(175) SACHS, E. "Lesions of the frontal lobe: A review of 45 cases." *Arch. Neurol. and Psychiat.,* 24: 1930, 735–742.
(176) SACHS, E. "Symptomatology of a group of frontal lobe lesions." *Brain,* 50: 1927, 474–479.
(177) SARGANT, W., and BLACKBURN, J. M. "The effect of benzedrine on intelligence scores." *Lancet,* 231: 1936, 1385–1387.
(178) SCAMMON, R. E. "The literature on the growth and physical development

of the fetus, infant, and child: A quantitative summary." *Anat. Rec.*, 35: 1927, 241–267.

(179) SCHWAB, S. "Changes in personality in tumours of the frontal lobe." *Brain*, 50: 1927, 480–487.

(180) SHERLOCK, E. B. *The Feeble-Minded: A Guide to Study and Practice.* (Macmillan: London, 1911) 327 pp.

(181) SHERRINGTON, SIR C. *The Brain and Its Mechanism.* (University Press: Cambridge, 1937) 35 pp.

(182) SHOCK, N. W. "Some psychophysiological relations." *Psychol. Bull.*, 36: 1939, 447–476.

(183) SHOCK, N. W. "Physiological factors in mental development." [To appear in] *Review of Educational Research.*

(184) SMITH, J. R. "The electroencephalogram during infancy and childhood." *Proc. Soc. Exper. Biol. and Med.*, 36: 1937, 384–386.

(185) SMITH, J. R. "The electroencephalogram during normal infancy and childhood: I, II, and III." *Ped. Sem. and Jour. Genet. Psychol.*, 53: 1938, 431–482.

(186) SMITH, J. R. "The 'occipital' and 'precentral' alpha rhythms during the first two years." *Jour. Psychol.*, 7: 1939, 223–226.

(187) SOUTHERLAND, R. W. "Three cases of tumors in posterior cranial fossa with mental symptoms." *Bull. Neurol. Inst. N. Y.*, 2: 1932, 144–156.

(188) SPELT, D. K. "Conditioned responses in the human fetus in utero." *Psychol. Bull.*, 35: 1938, 712–713.

(189) STONE, C. P. "The effects of cerebral destruction on the sexual behavior of male rabbits. III. The frontal, parietal and occipital regions." *Jour. Comp. Psychol.*, 6: 1926, 435–448.

(190) STRAUCH, F. W. "Die Beurteilung geistiger Erschöpfung im Schulalter; Ein Beitrag zur Phosphorsäurebehandlung." *Medizinische Klinik*, 19: 1923, 209.

(191) STRAUSS, M. "The spastic in our schools." *Crippled Child*, 15: 1938, 160–163.

(192) STRAUSS, I., and KESCHNER, M. "Mental symptoms in cases of tumor of the frontal lobes." *Arch. Neurol. and Psychiat.*, 33: 1935, 986–1007.

(193) SUTTON, H. "The influence of high temperatures on the human body, especially with regard to heat-stroke." *Jour. Pathol. and Bacteriol.*, 13: 1909, 62–73.

(194) TILNEY, F., and CASAMAJOR, L. "Myelinogeny as applied to the study of behavior." *Arch. Neurol. and Psychiat.*, 12: 1924, 1–66.

(195) TILNEY, F., and RILEY, H. A. *The Form and Functions of the Central Nervous System. An Introduction to the Study of Nervous Diseases.* (Hoeber: New York, 1923) 1019 pp.

(196) TOWN, C. H. *Familial Feeble-Mindedness.* (Foster and Stewart: Buffalo, 1939) pp. 1–97.

(197) TREDGOLD, A. F. *Mental Deficiency (Amentia).* (William Wood: New York, 1922) 4th ed., 569 pp.; (6th ed., 1937) 556 pp.

(198) TRYON, R. C. "Studies in individual differences in maze ability. I. Measurement of the reliability of individual differences." *Jour. Comp. Psychol.*, 11: 1930, 145–170.

(199) TRYON, R. C. "Studies in individual differences in maze ability. II. De-

termination of individual differences by age, weight, sex and pigmentation." *Jour. Comp. Psychol.,* 12: 1931, 1–22.

(200) TRYON, R. C. "Studies in individual differences in maze ability. III. Community of function between two abilities." *Jour. Comp. Psychol.,* 12: 1931, 95–115.

(201) TRYON, R. C. "Studies in individual differences in maze ability. IV. The constancy of individual differences: Correlation between learning and relearning." *Jour. Comp. Psychol.,* 12: 1931, 303–345.

(202) TRYON, R. C. "Studies in individual differences in maze ability. V. Luminosity and visual acuity as systematic causes of individual differences, and an hypothesis of maze ability." *Jour. Comp. Psychol.,* 12: 1931, 401–420.

(203) TRYON, R. C. *The Genetics of Learning Ability in Rats — A Preliminary Report.* (University of California Press: Berkeley, 1929) University of California Publications in Psychology, Vol. 4, No. 5, 71–89.

(204) TRYON, R. C. "Individual differences." [In] Moss, F. A., editor: *Comparative Psychology.* (Prentice-Hall: New York, 1934) 529 pp.

(205) TRYON, R. C., TOLMAN, E. C., and JEFFRESS, L. A. "A self-recording maze with an automatic delivery table." *Univ. Calif. Publ. Psychol.,* 4: 1929, 99–112.

(206) WARDEN, C. J., JENKINS, T. N., and WARNER, L. H. *Comparative Psychology.* (Ronald Press: New York, 1935 and 1936) Vol. 1, "Principles and Methods." 506 pp.; Vol. 2, "Vertebrates," 560 pp.

(207) WEISENBERG, T., and McBRIDE, K. E. *Aphasia: A Clinical and Psychological Study.* (The Commonwealth Fund: New York, 1935) 634 pp.

(208) WILLIAMSON, R. T. "On the symptomatology of gross lesions (tumors and abscesses) involving the prae-frontal region of the brain." *Brain,* 19: 1896, 346–365.

(209) YERKES, R. M. "The intelligence of earthworms." *Jour. Animal Behavior,* 2: 1912, 332–352.

SECTION IV

THE PROBLEMS OF NATURE-NURTURE RESEARCH: THE RELATIONS BETWEEN INTELLIGENCE AND SUCH FACTORS AS SEX, RACE, PERSONALITY, SOCIO-ECONOMIC STATUS, AND THE ENVIRONMENT IN GENERAL

CHAPTER V

INTELLIGENCE AS RELATED TO SOCIO-ECONOMIC FACTORS

JANE LOEVINGER
Institute of Child Welfare
University of California
Berkeley, California

I. ANALYSIS OF THE PROBLEM

Questions concerning the influence of hereditary and environmental factors upon intelligence are as old as the science of individual differences. Every fresh attempt to attack the problem experimentally has revealed new complexities; taken together, the various approaches show that the results obtained are to an important extent dependent upon the method of analysis. When broad studies are replaced by investigations confined to specific aspects of the environment, like socio-economic differences or rural-urban differences, the problem is hardly simplified, for we then have the additional task of separating the effects of these factors from those of other environmental factors. Here, then, methodology becomes even more crucial in determining results. The following review of evidence has consequently been organized chiefly with considerations of method in mind.

To make a statement about the influence of socio-economic status on intelligence one must ascertain two sets of facts: (1) What is the degree of the relation between socio-economic status and intelligence? (2) What is the nature of the relation? The first of these problems, the magnitude of relation, can be resolved into two: (*a*) How large are the differences between social classes in relation to the differences within classes? (*b*) To what extent can results obtained with a given set of measures and a given sample be generalized to other measures and other samples? The second problem, the nature of the relation, resolves into at least three issues: (*a*) distinction between the effects of social status on intelligence and influence of intelligence on social level attained; (*b*) interaction between intelligence, social level, and other variables in the populations studied; and (*c*) the aspect of social status that is responsible for whatever influence it has on intellectual development. These divisions of the problem are, it will be seen, not sharply separable. Before surveying the results in each of these categories, it will be well to equip ourselves with some methodological reservations.

II. Problems of Method

1. Considerations in Determining the Degree of Relation

a. Measures of Intelligence. The first authors to point out the relation between intelligence and social level measured intelligence in terms of such crude estimates as could be obtained from gross achievement and biographical data. Later reports have dealt with results from standardized tests. During the period covered by this review, the tests employed have differed widely in administration and content. As to administration, both individual and group tests have been used. As to content, use has been made of verbal tests, non-verbal performance tests, and composite tests of the Stanford-Binet variety. The composite tests are generally conceded to emphasize verbal factors, although this does not obtain equally at all age levels. Even the most verbal of the composite tests at the preschool level utilize more performance material than is found in the same scales at the higher levels. In all infant tests, used below about eighteen months, non-verbal content is, of course, predominant.

Little question has been raised about the equivalence of group and individual tests for purposes of comparison of social classes, probably because similar results have been obtained with the two types of measures.

Performance tests, on the other hand, have almost uniformly yielded less marked socio-economic differences than verbal tests at corresponding age levels. This has frequently been interpreted as meaning that 'intelligence' is not really affected by the differences in socio-economic status, even though verbal measures of intelligence are so affected, or alternatively, that verbal measures of intelligence are 'unfair' to persons raised in a culturally poorer environment. If, however, the correlation between verbal and performance tests is less than perfect, then groups selected for differences in verbal intelligence will ordinarily show regression toward the mean of the population with respect to performance tests. Such a result is not in itself a substantial demonstration of the unfairness of verbal tests to the groups involved.

Where the users of infant scales have found less marked differences between social classes than those obtained with verbal tests on the higher age levels, they have interpreted their results to mean either that the environment had not yet had a chance to work its beneficent or deleterious effects, or that the same functions were not being tested with the different tests. An argument essentially similar to the former interpretation reasons that the divergence with age of mental-growth curves for the different classes proves that the effects of environment are marked and cumulative. The obvious difference in content of infant scales and scales for older children should, however, induce caution in comparisons involving the two types of tests. It is, furthermore, a fact to be demonstrated, not assumed, that where age changes in intelligence are measured by a single test, the same functions are involved at the different ages. At least one study (53) [1] has suggested a systematic change in the content of the 1917 Stanford-Binet from one age to the next.

The foregoing discussion has dealt essentially with two points. The first is formulated by Anastasi (1) in the following terms:

> There is no substitute for verbal tests. It is a psychological impossibility to alter the verbal content of a test without altering the mental processes involved. (p. 483)

The second point is that there is no proof that one can alter the difficulty of a test very much without altering the mental processes involved (45).

Even within the narrow age range found in a relatively select group of first-grade children, Sangren (75) has shown wide differences both in mean mental ages obtained with eight different tests and in the intercorrelations of the tests. One of his most striking findings is that the degree of validity assigned to a test varies markedly with the method of validation used.

It is not meant, however, to suggest that no means exist for ascertaining the differential effect of socio-economic factors on different mental functions. A study to determine such influences should be based on a relatively narrow age range and should utilize tests of approximately equal difficulty and equal validity, where validity means ability to separate groups known to differ grossly in average intelligence. Under these conditions, variability in the amount of relationship to environmental factors becomes significant indeed.

b. Measures of Socio-Economic Status. In view of the preponderance of reports based on the occupation of the father, we may con-

[1] See the numbered references at the end of this chapter.

to age may have influenced the correlations. It is not clear whether the correlation with the Binet test means that the environmental factors tested are related to intelligence or that the Apperception Test is itself directly affected by intelligence. This test has, however, the virtue of being the only measure of socio-economic status the purpose of which is not immediately apparent. It is thus more difficult to falsify answers in order to produce a desired impression.

Composite scales, utilizing both occupational and non-occupational information, have also been developed to measure socio-economic status. Among the earliest of these is the Chapman-Sims Scale (23), which, like its successor, the Sims Score Card for Socio-Economic Status (81), is a questionnaire concerning the cultural, economic, educational, and occupational status of the family, to be filled out by the child. The questions included on the Sims Score Card were chosen from a larger battery on the basis of their efficiency in differentiating pupils in schools selected for differences in social status. Sims found a reliability of .95 $\pm$ .01 for two groups of 100 paired siblings, but similar results have not been uniformly obtained by later investigators (91).

The University of California Socio-Economic Index (6) consists of a weighted average of the following factors: family income transformed into logarithmic scale scores; years of schooling of the father; years of schooling of the mother; a Taussig rating of the father's occupation; and a composite rating of home, living room, and neighborhood. Bayley and Jones (6) found these factors to show intercorrelations ranging from .24 to .80.

Probably the most carefully validated and standardized measure is Leahy's (60) Minnesota Home Status Index, which yields a 'home status profile,' consisting of sigma scores on six indices relating to children's facilities, economic status, cultural status, sociality, occupational status, and educational status. Each of the indices, except the last two, is based on from eleven to thirteen questions asked of one of the parents. The occupational status index consists of a score between one and eight, depending on which category the father's occupation falls in, using the Minnesota Scale for Occupational Classification. Educational status is a combined rating for mother's and father's education. The index was standardized on a group of over 600 urban families in Minnesota, similar to the urban population of Minneapolis in terms of father's occupation and number of children. The questions included were selected by the method of internal consistency from more than eighty items in the original tentative scale, and they were assigned to the different indices by expert judges. Uncorrected intercorrelations of the indices range from .47 to .73. Economic status, cultural status, sociality, and children's facilities have split-half reliabilities ranging from .82 to .67, in the order named. When the intercorrelations of these four indices are corrected for attenuation, they range from .82 to .90. For 200 cases, the correlation of the total index with the Sims Score Card was .94.

A more complete review of these and other measures of social status will be found in Leahy's monograph (60). Some characteristics of the scales are already apparent. (1) They involve variously occupation, education, income, cultural possessions, other home and neighborhood conditions, and combinations of measures based on these factors. (2) Most of them assign some number or score to one's socio-economic status. The fact that in studies of composite scales intercorrelations between factors range from .24 to .90, with most of the correlations falling around .6, suggests that the different indices may not all be measuring the same thing. On logical grounds we may question the complete measurability of social status. Let us consider the two points separately.

From one point of view, it may be suggested that the very attempt to construct a scale of socio-economic status involves a 'quantitative fallacy'; namely, the notion that all differences are susceptible of linear quantification, which has so stimulated American psychology and has led it into so many premature generalizations. In measuring socio-economic status we tacitly assume that it is a linear variable, that for every pair of individuals a meaningful 'more than' or 'less than' statement about their relative statuses can be made. To the extent that this assumption is correct, we may hope to create a valid measure of socio-economic status. To the extent that the assumption is incorrect, our measurement is defective. Errors due to the inaccuracy of assumptions are easily neglected, for they cannot be estimated as precisely as errors of measurement and errors of sampling. So much the more imperative, then, that we examine the appropriateness of the 'linear quantity' assumption with respect to scales of socio-economic status.

While it is clear enough that the socio-economic status of the owner of a mill is superior to that of his mill hand, it is less clear as to what relative position shall be assigned to their children's school teacher. If the company lawyer has an inferior economic status to the owner of the mill because he is in the employ of the mill, is the union lawyer in a similarly inferior position with respect to the mill hands? Actually, the psychologists and economists who have constructed some of the most commonly used scales have rated themselves superior in socio-economic status to the owners of large businesses, though clearly they are inferior in economic status and perhaps in the eyes of the majority of the people they are inferior also in social status. Intellectuality of occupation, which is the basis for many of these scales, does not bear a one-to-one relation either to social or to economic status. Goodenough[1] has recognized clearly the non-quantitative aspect of socio-economic differentiation, and cautions against the use of occupational classifications as scales or measuring devices. Many investigators do, nonetheless, treat occupational categories as a linear continuum, and the Barr Scale does so in its very construction. All of the non-occupational and composite measures of socio-economic status are continuous scales on which an individual's position is represented by a single number or at most by several numbers.

The wide range of intercorrelations between socio-economic factors has already been cited as evidence against the univocity of the concept. Somewhat similarly, a recent study by Davidson and Anderson (31) criticizes the 'economic' part of the concept of socio-economic status on the basis of data from the community of San Jose, California. Using a scale that was essentially the same as the various Taussig modifications, they found a large range of incomes in each of the six categories with considerable overlap between categories. During the years 1928 and 1929 the mean annual income of workers in the second category (proprietors, managers, and officials) was greater than the mean income of those in the first or professional group; since 1929 the relationship has been reversed. The same writers criticize the 'social' part of the term 'socio-economic status' because of lack of evidence lending concrete meaning to it.

How, then, can we rescue meaning for the concept of socio-economic status? It is pertinent to inquire whether the components that any scale shares with intelligence are also shared with other differently constructed socio-economic scales. If it should appear that, for a set of measures such as those discussed above, the components that any one of them has in common with the child's intelligence also enter into the other measures of socio-economic status, a rationale is provided for using the common term 'socio-economic status' for each of these scales. Research could be set up to investigate this point by multiple factor

[1] See Chapter X.

analysis. In lieu of such evidence, the term 'socio-economic status,' or, for short, 'social status,' will be used in this review much as it has been used in the original studies, with the understanding that this represents convenience and may not be fully justified technically. Similarly, the phrase 'social classes' will be used, not to designate classes in the sense of any economic theory, but merely as a collective term for the classes defined by a given scale of socio-economic status.

c. Quantitative Statement of Relationship. The most common method of expressing the degree of relationship between the two variables has been comparison of group averages. The meaning of a given group difference may in turn be stated in terms of 'statistical significance' or 'level of significance.' Statistical significance refers to the probability that the obtained difference holds true for the populations of which the tested groups are random samples. When the probability that the obtained difference does not hold true for the larger populations becomes less than five chances in 100, we say the difference passes the 5 percent level of significance. Similarly, when there is only one chance in 100 that the obtained difference does not hold true for the whole populations, we say the difference reaches the 1 percent level of significance. By increasing the size of the sample, any real difference between the means of two groups, however negligible in comparison with intra-group or individual differences, can be made statistically significant. The statistical significance of a set of data is indeed important, and ought to be computed more often than it is, but we must not be so greatly impressed by statistical significance that we omit consideration of psychological and practical significance.

The 'practical' significance of the results depends on how great the differences between averages for different social classes are in comparison with the total amount of variation in the population when all classes are considered together. The correlation coefficient expresses the amount of variation that two variables have in common in terms of the total amount of variation in the two variables. Neff (67) objects to the use of the correlation coefficient on the grounds that the distribution of intelligence is normal while the distribution of socio-economic status is markedly skewed. Even granting that neither in the case of intelligence nor social status are we possessed of anything like absolute units of measurement, the difference between the forms of the distributions is sufficiently independent of the specific instruments employed to suggest that it is a substantial fact. The fact that the correlation between the two measures cannot reach unity, it might be argued, accurately reflects the fact that prediction from a very skewed to a normally distributed variable cannot be perfect.

Two methods have been used to overcome Neff's objection. One is the computation of correlation ratios (*eta*) rather than Pearson product-moment coefficients. Correlation ratios do not assume linear regression. The second method has involved rescaling the socio-economic index. For example, in computing the University of California Socio-Economic Index, the logarithm of the income is used rather than the absolute value (6). This usage is justified by the reasoning that equal increments of income have different significance as parts of the child's environment, depending on the original magnitude of the income. The difference between an income of $1,000 and $2,000 will affect the child more than a change from $10,000 to $11,000.

In two studies to be reported (20, 65) correlations have been computed between average values for groups rather than between two sets of scores on individuals. It will be apparent that the magnitude of the correlations based on groups is not comparable to corresponding correlations based on individuals.

2. Considerations Involved in Analyzing the Nature of the Relationship

a. Hypotheses as to Cause and Effect. Granted that there exists a correspondence, more or less close, between the average intelligence of a social group and its place in the social scale, to what factor shall we ascribe causal importance? A correlation between two variables does not tell us in the first instance whether one of the two variables or still others are causally responsible. The most common arguments to account for the correlation between socio-economic status and intelligence concern only the direction of influence between the two major variables, but they are by no means the only permissible explanations of the facts.

One common hypothesis is that the superior advantages of a home with high cultural standards tend to aid or promote mental development. A second frequently defended hypothesis is that persons of superior intelligence gravitate toward the occupations that stand higher on the scale; on the average they earn more money to give their children more cultural advantages, and they transmit superior abilities to their children by way of biological heredity. There are no longer any investigators who voice serious doubts as to the efficacy of either of these factors, at least in extreme cases. Some of the disagreements about the relative efficacy of the two factors will be found to disappear when results are viewed in relation to the stated levels and ranges of ability and social status for which they have been shown to hold. An attempt in that direction is made in this review.

A third hypothesis, logically quite as valid as the foregoing two, has not been stressed often. The relation between socio-economic status and intelligence may be attributed to the fact that both are related in the populations studied to other variables, such as race, personality factors, family size, and probably other unidentified factors. Where a third variable or set of variables can be shown to be related to the two major variables, it does not follow immediately that it is responsible for a spurious relation between intelligence and social status. The direction of causal influence is no easier to establish here than when only intelligence and socio-economic status are considered. In some instances of the association with third factors it will seem easier to conceive of social status as responsible for a spurious relation with the other factor and intelligence; in other cases it may be easier to conceive of intelligence or of the third factor as responsible for a spurious relation between the other two. Because of the absence of objective evidence, no such conclusions will be drawn in this review.

One further question as to the nature of causal influence concerns the mediating aspect of socio-economic status. In a sense it seems strange that anyone should suppose that so abstract an attribute as one's socio-economic status can influence anything at all. Clearly it is not socio-economic status as such but some aspect of it or concomitant of it that influences the development of intelligence. Under what condition, then, is it justifiable to speak of the influence of socio-economic status on intelligence? By analogy, we should not speak of the influence of environment on intelligence if it could be shown that the only measure of intelligence so affected was, for example, memory span for digits. If, however, each, or a majority, of the subtests of the Stanford-Binet test was shown to be influenced by environment, it would be justifiable to speak of the influence of environment on Stanford-Binet intelligence as a whole. Similarly, if a number of relatively independent measures of social status can each be shown to be related to intelligence, there is excuse for describing the relation in terms of social status instead of in terms of a specific measure of social status. The possibility remains that each of the correlations is due to the association of intelligence with still another factor that happens, in that population, to be related to the various measures of social status. This is less probable if the correlation of intelligence with a composite of the various aspects of social status is higher than with any one of them by more than the increase in reliability of measurement would lead one to expect. Few studies can claim to have presented crucial evidence on these questions.

b. Sources of Evidence. Some studies (5, 26) have presented evidence on the tendency for superior persons to enter the more intellectual occupations. These studies will not be reviewed, as they have not been crucially conceived. They have not controlled the possibility that children's occupational choice as well as their intelligence may be affected by parent's occupation. To some extent that must be the case.

Evidence in favor of production of intellectual differences by differences in socio-economic status independent of hereditary differences has come from study of a very few identical twins reared apart from infancy. Crucial evidence in favor of production of intellectual similarity by socio-economic similarity, independent of genetic similarity, has come especially from studies of foster children. To the extent that the above similarities and differences do not equal corresponding ones in unselected populations, such studies become evidence for the influence of heredity and the joint influence of heredity and environment.

Certain studies on the intelligence of children living in institutions for varying lengths of time have been interpreted as relevant to the subject of socio-economic status and intelligence, on the grounds that the institution represented a lower socio-economic level than that for some comparison group, or because it represented a higher level than the homes from which the children were taken, or because it represented a relatively homogeneous social environment. While these studies are of interest, in no case can they be unambiguously interpreted with respect to the problem entertained here.

Still another line of evidence is represented by Van Alstyne's (93) comparison of correlations of various environmental factors with the parents' and the child's intelligence. The reasoning, which she admits is a bit risky, is that if the environmental factors correlate equally with the intelligence of mother and child, and cannot have caused the intellectual level of the mother, it is not necessary to say that they caused the intellectual level of the child.

c. Statistical Techniques. Statistical techniques for analyzing the nature of the relationship have not changed since the publication of the *Twenty-Seventh Yearbook* of this Society. It is still necessary to take the precautions in interpretation suggested by Burks and Kelley (15) in that Yearbook. A few points will be added to the discussion.

Where partial correlations have been computed by the authors of the works reported here, they will be neglected, since nowhere are they accompanied by demonstration that the assumptions underlying the statistics are fulfilled. Probably in every case it could be shown, on the grounds that Burks discusses, that partial correlation is inappropriate.

Burks' objection that partial correlation partials out too much applies just as well to 'holding a factor constant' experimentally as to partialling the factor out statistically. Holding constant a third factor, such as race, results in holding constant all those factors that vary with race, just to the extent of their covariance. Furthermore, holding race constant may result in varying systematically an unknown number of sources of variation in intelligence and social status that would normally vary in an unsystematic fashion and thus influence the discovered correlation very little. Since these effects take place in the subterranean recesses of 'experimental error,' we have no measure of the influence on discovered relations.

A great deal of discussion has centered around the analysis by the method of path coefficients of the proportional contributions of environment and heredity to IQ variance in several of the samples reported below. While these discussions are worth studying both for methods and for information, the reader is warned to consider them with the following reservations: (1) The contribution of parental intelligence means only the contribution as judged from the estimate of intelligence yielded by the administration of intelligence tests to adults. (2) Environment means only the aspects of environment measured by the particular, fallible instruments employed. (3) No measure is obtained of the extent to which apparent contributions of the measured variables come in reality from other associated variables. (4) Complicated statistical techniques are based on assumptions that are never more than approximately satisfied, and the results are subject not only to errors of sampling and errors of measurement but also to errors due to only approximate satisfaction of assumptions. (5) Finally, the exact percentages of determination are based on samples (identical twins and foster children) that have no claim to representing the population at large.

The interacting associations of socio-economic status and intelligence with other variables in the populations studied suggest the use of a statistical technique that has not yet been systematically tried in the field; namely, analysis of variance. A set of several factors, each of which promised to show significant relation with intelligence — such as occupational status, racial background, parents' education, and so forth — could be selected. In order to use analysis of variance, it would be necessary to pair each value of each factor, or 'dimension,' with each value of every other factor. It will be seen immediately that this involves restriction of the range available for each factor. More specifically, if one were to use mid-parent education and occupational status as two of the factors, dividing occupational status into seven groups and mid-parent education into five groups or ranges, it would be necessary to find individuals in each of the seven occupational groups belonging in each of the five classes with respect to mid-parent education. The value of this approach would be in permitting an estimate of the relative proportions of the total variation in intelligence score ascribable to each of the factors studied, within whatever range it varied in the experimental set-up. The limitations of the approach would be the fact that only a few such factors can easily be studied at once, and that the factors can vary only within rather narrow ranges if we are to deal with a representative group for each category. In view of these difficulties, and the associated, perhaps insurmountable, sampling problems, it remains for future investigators to decide whether the information yielded would make it worth while to attempt such a study.

III. A Survey of the Evidence

Any survey of literature in a broad field involves selection; it is only fair to the reader to state the principles of selection. This review covers almost exclusively research appearing after or shortly before the *Twenty-Seventh Yearbook* of this Society. Prior studies are covered in earlier reviews (11, 63, 66, 77, 86). In a few cases older investigations have been included here, either as classics or as better illustrations of a point than subsequent work. Selection within more recent work has been based on degree of generality of interpretation possible and crucial isolations of factors studied. By 'degree of generality' is meant the extent to which the sample tested can be thought of as representing a wider and defined population; crucial isolation of factors has been discussed above.

Most of the data reviewed will be reported to fewer decimal places than in the originals; longer numbers are harder to read and less likely to be remembered, and the magnitude of the errors of sampling and measurement usually makes the seemingly greater accuracy more apparent than real.

1. Studies Concerned Primarily with the Degree of Relationship

a. Representative Populations of Different Ages. We might classify the studies of the magnitude of the relation in terms of the measure of intelligence used, or the measure of socio-economic status used, or in terms of the sampling. The last criterion is most convenient, since a number of the studies have employed more than one measure of intelligence or of socio-economic status. We shall consider first data that have some claim to being representative of the 'generality,' in order of chronological age, beginning with infancy. Discussion of groups selected for intellectual deviation, the feebleminded and the gifted, will follow in subdivision *b*, below.

(1) Furfey. Furfey (36) studied 45 to 49 infants at each of ages 1, 2, 4, 6, 9, and 12 months, taken in order of birth at a large city hospital. He found no significant relation between the score on the Linfert-Hierholzer Infant Scale and the parents' score on the Chapman-Sims Scale of socio-economic status. Eighty of those tested at 6, 9, and 12 months were retested at an average age of 4 years, 8 months (standard deviation, two months) with the Stanford-Binet (37). The correlation between the Linfert-Hierholzer scores and the Stanford-Binet scores was low, unreliable, and negative. The algebraic difference between Binet IQ and Linfert-Hierholzer IQ correlated .33 ± .07 with the Chapman-Sims score, indicating merely that the Stanford-Binet is related to the Chapman-Sims test at the higher age level.

(2) Dubnoff. In reviewing a study of Russian children perhaps one should refrain from using the term 'socio-economic status.' Dubnoff (32), using family income as an environmental measure, obtained results different from those of Furfey and those of Bayley and Jones reported below.

Three samples of children were tested with the California First-Year Mental Scale, at ages 3, 6, and 10 months; the correlations with family income were, at the successive ages, .27, .44, and .44, with P.E. approximately .07. Dubnoff's tentative interpretation of the findings emphasizes selective or genetic factors:

> It might be supposed that in the Soviet Union the attempted equalization of medical care and other environmental factors which are of potential importance in child development would tend to hold down any environmental correlations. On the other hand, it is the impression of the writer that because of greater opportunities for individual advancement, income is in the Soviet Union a better measure of individual differences and of ability than in other countries. It is possible that the positive correla-

tion with economic status, found in this sample at a considerably earlier age than in studies in the United States, reflects not an environmental factor in child development but an efficient selective factor in the economic system, and a parent-child correlation based chiefly on genetic factors in a population of very wide variability. (p. 72)

(3) Goodenough. One of the earliest and most representative groups of preschool children was that tested by Goodenough (40, 41), including 380 children between the ages 18 and 54 months, carefully matched for father's profession against the census proportions for Minneapolis. Her measure of intelligence was the Kuhlmann Revision of the Binet tests, administered twice, about six weeks apart. While the results of the second test are not comparable with results reported by other investigators, they are worth including, since by several criteria Goodenough found them to be more valid than the results with the first test. Her measures of socio-economic status were father's occupation, classified according to the Minnesota Scale for Occupational Intelligence, and number of years of education for the father and mother. Mean IQ's of children whose fathers are in each of the professional groups are shown in Table I. She found approximately the same results for children aged 2, 3, and 4 considered separately. The three possible types of explanation for the appearance of occupational differences in the intelligence scores of preschool children are, she points out, (1) that social transmission accounts for the occupational differences, (2) that industrial selection coupled with biological heredity is responsible, and (3) that the choice of tests is such as to maximize the influence of training and opportunity. For 213 of these cases information on the education of both parents was available. Most of the cases omitted for lack of information were from the lower social classes. For the smaller group, with restricted range of intelligence and social status, the correlation between child's IQ on first test and mother's education was .32. The correlation with father's education was .26. The corresponding correlations with the second IQ were both .35; probable errors were of the order of ± .04.

(4) Coffey and Wellman. These investigators (27) studied 417 children in the preschool laboratories of the Iowa Child Welfare Research Station between the years 1921 and 1934, tested approximately every six months. The measures of intelligence were the Stanford-Binet for children over 3 years, 6 months and occasionally for younger children; the Kuhlmann-Binet for the rest of the younger children. As the mean IQ's on the two scales for three-year-olds being tested for the first time was the same, the results were not treated separately. Socio-economic status was measured by means of father's occupation (classified in the same categories as in the study by Goodenough reported above) and number of years the parents spent in school. The average IQ's on entering preschool are shown in Table I for the different occupational classes. The only significant difference is that between Group I and

Groups II, III, and IV combined. While the IQ gain on first retest averaged between 5 and 7 points, there were no significant differences in amount of gain between ages, sexes, or occupational groups. Within each occupational group the amount of gain was related to initial score, with the lowest gaining most, though this tendency is probably an artifact caused by statistical regression. The comparison based on mid-parent education showed that children whose parents were better educated had higher initial scores but did not gain more. It also showed that the majority of children in the study came from families where the mid-parent education was 16 years of schooling or more. The group can scarcely be considered as representative of the population at large, and the nature of the sampling within the occupational categories is not easily calculable.

TABLE I. — MEAN IQ'S OF PRESCHOOL CHILDREN, CLASSIFIED BY FATHER'S OCCUPATION *

Father's Occupation	*Goodenough (41)*			*Coffey and Wellman (27)*	
	Number	*First Test*	*Second Test*	*Number*	*At Entrance*
I. Professional	56	116	125	278	119
II. Semiprofessional and managerial	29	112	120	47	110
III. Clerical and skilled trades	129	108	113	57	114
IV. Semiskilled and minor clerical	79	105	108	35	111
V. Slightly skilled	48	104	107	0	
VI. Unskilled	39	96	96	0	

* Adapted from Goodenough (41, p. 287) and Coffey and Wellman (27, p. 192).

(5) Bayley and Jones. A connecting link between the findings of Furfey and of Goodenough is furnished by the report of Bayley and Jones (6) on an intensive longitudinal study of children born in Berkeley in 1928 and 1929. The sample includes 61 children from 59 families, heterogeneous with respect to social status but somewhat superior to the community average. The measures of intelligence were the California First-Year Mental Scale during the first year, the California Preschool Schedules I and II administered twelve times from 18 to 60 months, and the Stanford-Binet at 72 months. At 66 months a vocabulary test from the Thorndike CAVD test and a series of form boards were administered. Ratings of the family on the California Socio-Economic Index were obtained. Correlations between the environmental variables making up the index and intelligence-test scores tend to be negative in the first 6 months, hover around zero up to about 18 months, and are increasingly positive as age increases. From the age of 21 to 72 months, the various environmental factors maintain quite consistently the following order with respect to closeness of association with intelligence: mother's education, father's

education, father's occupation, social rating (home and neighborhood), and family income. The total socio-economic rating fluctuates in its degree of relation to intelligence; it tends to be related somewhat more closely than the social rating scale but less than father's occupation. As representative of the findings, we may take the correlation between the various factors found at the ages of 30 and 72 months. There were 41 cases at 30 months, 42 cases at 72 months. The correlations of intelligence-test scores with the environmental factors were, respectively, mother's education, .48 and .58; father's occupation, .25 and .38; total socio-economic scale, .20 and .41; social rating, .19 and .29; family income, .08 and .32. Among the other findings are that environmental variables show lower correlations with scores on motor-ability tests and that the vocabulary and form-board scores correlate about .4 with mid-parent education.

(6) Standardization Group of Terman and Merrill. For representativeness of American children, no sample has equalled that assembled by Terman, Merrill, and their co-workers (91) in the standardization of the revised Stanford-Binet, and the feat is not likely to be duplicated in the near future. For that reason their results are reported in some detail in Table II. The IQ estimate is a composite of the scores on Form L and Form M for each subject. The measure of social status is father's occupation classified according to the Minnesota Scale for Occupational Classification. None of the age changes is consistent through all levels, with the exception of those of the children in the lowest occupational group. A comparison of the scores in the various occupational categories with those in corresponding categories for Goodenough's study of preschool children testifies to her success in getting a representative sampling of city children.

(7) Duff and Thomson. A widely quoted early study on the relation of father's occupation to intelligence is that of Duff and Thomson (34) who gave the Northumberland Mental Test to all elementary-school children in Northumberland (except in Newcastle and Tynemouth) who were over 11 and under 13 years of age, and to most secondary-school children between those ages. Parents' occupations were available for 13,419 of the subjects. The IQ's ranged from 112 for those whose fathers were in the professions (N = 137) to 96 for those whose fathers were laborers, or in 'low-grade occupations' (N = 1.214).

(8) Stoke. This investigator (85) studied 508 children living in a Boston suburb, all in the lower grades, all of North European stock. The number of siblings included in the study was small. He classified father's occupation according to Taussig's grouping. The number of cases in the lower occupational categories was too small for the group to be considered a representative urban sample. The IQ's were also thrown into a five-group classification: those above 119, between 110 and 119, between 90 and 110, between 80 and 89, and those below 80. The classification of individuals was based on both a Stanford-Binet test and a Dearborn Series IA test. Where the two differed

slightly, more weight was given to the Stanford; where they differed markedly, the case was not used. The proportion of children above average in intelligence increased consistently with social class; the proportion below average decreased consistently with social class. An analysis of the Stanford-Binet

TABLE II. — MEAN IQ'S OF THE STANFORD-BINET STANDARDIZATION GROUP CLASSIFIED BY FATHER'S OCCUPATION

Father's Occupation	*Mean IQ by Chronological Ages, Years*				*Unweighted Average* *
	2 to 5 1/2	*6 to 9*	*10 to 14*	*15 to 18*	
I. Professional	116	115	118	116	116
II. Semiprofessional and managerial	112	107	112	117	112
III. Clerical, skilled trades, retail business	108	105	107	110	107
IV. Rural owners	99	95	92	94	95
V. Semiskilled, minor clerical and business	104	105	103	107	105
VI. Slightly skilled	95	100	101	96	98
VII. Day laborers, rural and urban	94	96	97	98	96

* These averages were computed from the data of Table XII in *Measuring Intelligence* (91, p. 48), where the IQ's are reported to one more place accuracy than is reproduced here. Neff (67) reports the averages as representative of the findings, neglecting the fact that the number of cases in the different age groups differed widely.

by items showed that children from the superior social classes were quite consistent in their superiority, but that they had the least advantage in tests requiring rote memory. For purposes of comparison with other studies, Stoke calculated a Pearson coefficient of correlation, which was .30 ± .027; however, he points out the inappropriateness of product-moment correlation for his data. The coefficient of mean square contingency (C) between intelligence and social status was .357, with the maximum possible being .896.

(9) Jordan. Twice a year for three years Jordan (54) examined 1,247 school children living in a North Carolina mill town. His measure of socio-economic status was occupation, classified according to the Taussig divisions, with a sixth category added for farmers. Children of mill workers comprised 578 members of his sample; children of farmers, 104; all other specific occupations were represented by smaller numbers. The measures of intelligence were the Pintner-Cunningham Primary Mental Tests and Dearborn Group

Tests of Intelligence, Series I, for Grades I and II, and the National Intelligence Test, Forms A and B, for Grades III through VII. Jordan divided IQ's into the five categories used by Stoke, and compared the percentage of children from each occupational group falling in each IQ class with corresponding percentages for Stoke's Boston group. The main differences lay in the lower scores for those in the unskilled labor group in Jordan's study, and in the large numbers of those below an IQ of 80 for all occupational categories. Considering the children of mill workers alone, of the 243 given a language test in 1925, 16 percent had IQ's between 70 and 79, and 11 percent had IQ's below 70; of the 103 children of mill workers who took non-language tests the same year (that is, children in the first two grades), less than 9 percent had IQ's between 70 and 79 and 4 percent had IQ's below 70. The IQ's on the language test were closely similar to those obtained by another investigator in another community of mill workers. The decrease of IQ with age was more marked for mill workers' children than for other children in the study. For 181 children who took both the Dearborn and Form A of the National Intelligence Test, the mean IQ difference was 8.75 ± 1.43.[1] A mean difference in IQ of 10.6 ± 1.36 appeared for the 227 who took both the Pintner-Cunningham Test and the National Intelligence Test. There was an insignificant difference in intelligence for 288 who took both non-language tests. From all these facts Jordan concludes that there is " an impossible number of feeble-minded present among the children of mill workers " on language tests, with a " more reasonable number " on non-language tests. The bases for this conclusion are vulnerable at a number of points. Stoke's group was a poor one to choose for comparison, as there is no assurance that it was a representative urban sample. A number of Jordan's comparisons involve uncritical lumping of IQ's obtained with different tests. And the IQ change of 10 points, more or less, corresponds closely with the amount of regression MacMurray (64) found when he compared children, selected for brightness and dullness on the Stanford-Binet, for IQ's on the Pintner-Paterson Performance Scale.

(10) Hildreth. Another study characterized by low IQ's in the lower occupational classes is that made by Hildreth (46) on school children in a city located in the oil-industry section of Oklahoma. Her subjects were about 600 of an enrollment of 2,000 students. They included all those for whom the parent's occupation was on the school record, but the selective factors involved were not discussed. Using the Taussig classification, the occupational distribution of her subjects was as follows: Class I, the highest or professional class, 7 percent; Class II, 25 percent; Class III, 42 percent; Class IV, 11 percent; Class V, 15 percent. The median Stanford-Binet IQ for the total group was 97. For the five classes separately the median IQ's were I, 113; II, 108; III, 98; IV, 84; V, 76. It is difficult to see why there should

[1] The text of the report does not make it clear whether probable error or standard error is used.

be so few students in the two lowest categories and why their intelligence ratings should be so low. Possibly the subjects represent a biased selection of the school population. More probably, the type of work available in the lower socio-economic brackets for that region is such as to attract workers deficient in either intellectual endowment or opportunities.

(11) Paterson. In the standardization of the Minnesota Mechanical Ability Tests (69), the tests were found to be virtually independent of measures of socio-economic status, in terms of occupation, education, and cultural status.

(12) Stroud. In another southern study, Stroud (87) tested all the school children of one Georgia county and 300 in another county by means of the Pressey Classification Test, primary and intermediate forms. Complete data were available for 1,057 cases between the ages of five and eighteen, all of native-born white stock, mostly children of poor farmers. Stroud felt that the differences of economic status, modes of living, and educational opportunities were small within the group examined. He used tax assessment as a measure of economic status. IQ correlated with tax assessment to the extent of .25, which was practically the same as the mean correlation between raw score and tax assessment for each of the ages separately. There was a smaller, but statistically significant, negative relation between intelligence-test score and number of children in the family.

(13) Sirkin. Sirkin (82) measured home environment of another sample from the Soviet Union in terms of occupation and education of father and education of mother, putting each family into one of six categories on the basis of all three factors. His subjects were the elementary-school pupils of Charkow, 854 in the fourth grade, 850 in the fifth grade, and 682 in the sixth grade. Although the proportion of students from different social classes varied in the various schools, the curriculum was the same; the measure of intelligence used was total score on five tests: reading and comprehension, mixed sentences, opposites, similarities, and analogies. For over 600 pupils in a single grade the reliability was .96, and the test-retest correlation with a six months' interval was .89. There was a consistent and regular increase in score with environmental level, for each grade separately. In each case the children from the highest social level were approximately one standard deviation above the mean for their grade, while those from the lowest social level were about half a standard deviation below the mean for their grade. The correlation ratios for score against environmental level were .39, .43, and .40 for Grades IV, V, and VI, respectively. Sirkin states that environmental level, as he defined it, is a qualitative variable, and he has no grounds for transforming it into a quantitative one. Treating the categorical ranks as measures and computing Pearson correlations for the variables, however, produced in no case a change of coefficient greater than .01.

(14) Chauncey. In the junior high school at Stillwater, Oklahoma, Chauncey (25) obtained measures of intelligence, achievement, and social status for 113 eighth-grade and 130 ninth-grade students. The Sims Score

Card correlated .21 and .19 with the McCall Multimental Scale, .30 and .35 with the Stanford Achievement Test. The McCall and Stanford tests correlated .57 and .59 with each other. In each case, the first of the above correlations belongs to the eighth-grade pupils.

(15) Byrns and Henmon. Results for 100,820 high-school seniors, examined between 1929 and 1933 by the Wisconsin state testing program, are reported by Byrns and Henmon (16). They feel that high-school seniors constitute a group for which differences in intelligence due to unequal education and language handicap are minimized, but that selective factors will not have operated equally for all social groups. The measures of intelligence were the Ohio State University Psychological Test in 1929, the American Council Psychological Examination in 1930 and 1931, and the Henmon-Nelson Test of Mental Ability in 1932 and 1933. Scores on the different tests were made comparable by use of percentiles. Father's occupation was classed into one of the following groups, in order of the median intelligence of the students: professional, executive, business, clerical, miscellaneous, skilled workers, semiskilled, unclassified and unemployed, unskilled, farmers. Differences between groups were for the most part statistically significant and consistent over the five-year period. The correlation between intelligence of the student and the rank of the father's occupation was .18, indicating the large amount of variation within each occupational group. Distortion of the scale of intelligence by use of percentile scores probably lowered this correlation.

(16) Yerkes and Cattell. The notion that an 'occupational hierarchy of intelligence' exists dates its popularity from the report of the median Alpha and Beta scores for numerous occupations represented in the United States Army during the World War (97). More recently, R. B. Cattell (18) has reported norms for various occupations in England on Scale III of the Cattell Group Intelligence Tests. His sampling methods were too heterogeneous to permit a statement as to exactly what population is represented by the sample. Median IQ's for the occupations studied ranged from 151 to 78, beginning with secondary-school teachers and physicians and surgeons, and ending with hair-dressers, upholsterers, welders, and factory packers and sorters.

College students represent a group so selected for intelligence that one would expect the relation between socio-economic status and intelligence to be greatly diminished or possibly even reversed. Four studies on unselected college freshmen, however, show significant relation between intelligence and father's occupational status.

(17) Haught. The largest group is that studied by Haught (44), who accumulated data on 3,414 freshmen of the University of New Mexico between 1921 and 1936. Those entering between 1921 and 1927 were given the Army Alpha; those entering from 1928 to 1936 took the American Council Psychological Examination. The scores were made comparable by turning

them into percentiles based on New Mexico freshmen. The mean percentile score for the professional group was 60; for the business and clerical group, 57; for those in the skilled labor category, 50; semiskilled, 45; and unskilled, 35. All differences between groups exceeded the 1 percent level of significance, but the greatest difference, relative to its standard error, was between the two lowest groups. Applicability of the critical ratio test to differences between mean percentiles is questionable; in fact, computation of mean percentiles is itself not strictly justifiable in terms of the derivation of the mean as the most characteristic or 'best-fitting' value. Mean percentiles cannot safely be considered other than as inexact approximations.

(18) Glass. Among the 118 students in a required freshman course, Glass (39) found no children of persons in the unskilled labor class. The mean percentile score on the Thurstone Psychological Examination for College Freshman was 65 for those students whose fathers were from the professional class; 54 for those from the skilled labor class; 52 for those from the commercial classes; and 45 for those from agricultural classes.

(19) Canady. A similar occupational gradation was reported by Canady (17) among 441 Negro college freshmen at the West Virginia State College. Unlike the previous studies on college students, he found the occupations of fathers of Negro college students to resemble the occupational distribution of Negro workers in the state. For example, 78 percent of his cases came from the skilled or unskilled labor groups. Although no conclusion can be, or was, drawn from the occupational distribution alone, it suggests some interesting possibilities. Using the American Council Psychological Examination as his measure of intelligence, he found the occupational groups to stand in the order: professional, with a median score of 98; commercial, 95; artisan, 94; skilled labor, 88; and unskilled labor, 73. The median score for the total group was 84.5, and the middle 50 percent fell between scores of 58 and 119.

(20) Cuff. This investigator (30) tested 758 college freshmen with the American Council test and the Sims Score Card. College grades were obtained from the registrar. Socio-economic ratings correlated .24 with intelligence, .10 with scholarship. To illustrate the low predictive value of these correlations, Cuff computed coefficients of alienation (k), which were .97 and .995, respectively. These coefficients show that socio-economic status as measured by the Sims scale is very little better than chance as an indicator of the intelligence of this group of college freshmen.

b. Intellectually Deviate Populations. We turn next to studies of the relation between socio-economic status and very superior or very inferior intelligence.

(1) Terman. In his chapter on the "Racial and Social Origin" of his thousand gifted children Terman (89) arrives at the following, among other conclusions: The gifted children showed an excess of urban parentage, of Jewish

parentage, of native American parentage, and a deficiency of Latin and Negro ancestry, as compared with the population at large. In terms of the Taussig classification and Barr ratings, the parents of gifted children pursue higher occupations than unselected adults. The neighborhoods are slightly above average, the homes are above average on the Whittier Scale, the mean income is well above the average for the population. The mean amount of parents' schooling is about twice as many years as for the general population, and so on. Finally, " The data of this chapter offer considerable indirect evidence that the heredity of our gifted subjects is much superior to that of the average individual." One might add that there is considerable direct evidence that the socio-economic environment of the gifted children is superior.

(2) Duff. In a follow-up of the Northumberland survey, Duff (33) found that the children of very superior intelligence came from families in the higher occupational strata far more often than did control children drawn from the same schools. His data, unfortunately, came from questionnaires that many of the persons solicited failed to return.

(3) Paterson and Rundquist. An approach complementary to the study of gifted children is the study of the social origins of feebleminded persons. The occupational status of 823 residents and 516 applicants for admission to the Minnesota School for Feebleminded at Faribault were studied by Paterson and Rundquist (70). Occupations were classified according to the Minnesota Scale of Occupational Intelligence. It was assumed that " for the low levels involved in this study " the Stanford-Binet and the Kuhlmann-Binet results available on the different patients were comparable. They found that the fathers of their subjects came almost solely from the skilled, semiskilled, and unskilled labor groups, whereas the professional, business, clerical, and technical occupations contributed a small number relative to their numerical strength in the city of Minneapolis. IQ was negatively related to occupational level among the feebleminded; in fact, the occupational distribution of fathers of the lowest grade of feeblemindedness approximated that of the general population, whereas the higher grades of feeblemindedness were recruited almost exclusively from families in the semiskilled and unskilled categories. They interpret this to mean that the lowest grade, the idiots, included many pathological cases that could occur at any level of society. It is also true, as they note briefly, that wealthier people more often send their feebleminded children to private institutions, particularly if the feeblemindedness is of the higher grade type.

(4) Bradway. Data on 439 inmates of Vineland Training School were compared by Bradway (8) with the results of Paterson and Rundquist. The residents of Vineland represent a somewhat different selection, being both state and private wards. A routine diagnosis of primary or secondary etiology had been made on all of their subjects; the diagnosis of primary etiology was made for patients " for whom there is no evidence of post-natal causation and whose family background plausibly indicates hereditary transmission of

deficiency." For 200 cases the etiology was unknown or mixed. The distribution of parental occupations for the 116 patients of primary etiology resembled closely the distribution for subjects at Faribault undifferentiated as to etiology; that is, there were excess proportions in the lowest occupational categories. On the other hand, the distribution for 123 patients of secondary etiology was symmetrical, resembling that of the random sample of Minneapolis adults. To this extent her data support the hypothesis of Paterson and Rundquist. Bradway found, however, a negative relation between intelligence and paternal occupation, not only within the primary etiology group, but also, to a smaller extent, within the group of secondary etiology. She suggested as at least a partial explanation that selective factors operate differently at Vineland and at Faribault.

2. Studies Concerned Primarily with the Analysis of the Relationship

In the many pages that have been written about the relation of intelligence and socio-economic status, there has been curiously little explicit disagreement about the magnitude of the relationship. Perhaps this is because, when it is granted that a relation exists, those who like may conclude, to paraphrase St. Paul, that the powers that be are ordained of biological heredity. The same degree of relation permits others to conclude that intellectual differences in large measure result directly from environmental influences, of which cultural status is one. Such conclusions have been based on insufficient evidence or on unrigorous reasoning, or both. Some of the studies, it is true, present more crucial evidence on the nature of the relationship than others, and it is with these studies we shall be concerned in this section.

a. Relative Influence of Environmental and Selective Factors.

(1) Foster Children. More or less experimental isolation of the effects of social environment from the effects of biological heredity in producing the correlation of intelligence and social status appears with the publication of the studies of Burks (12) and Freeman, Holzinger, and Mitchell (35) in the *Twenty-Seventh Yearbook.* Nothing short of a complete chapter could deal adequately with the methods, results, and the interpretations to which these historic researches and subsequent studies of foster children have given rise. The results reproduced here are chosen for comparison with other studies and do not represent the full contributions of the researches. Proportional emphasis on different studies has been calculated, not to reveal the relative importance of the studies, but to avoid repetition of data reported elsewhere in this Yearbook. In selecting from the quantity of data on

foster children, we shall look not only for crucial isolation of variables but also for the ranges within which the results have been shown to hold. Each of the studies should be inspected, so far as possible, for these features; first, how the children were selected for adoption and for study; second, the range of intelligence and social status that obtained among true and foster parents; third, 'selective placement,' or the matching of the child to the foster home. Only with such facts in mind can we meaningfully compare the relation between socio-economic status and intelligence for foster and control children, or can we evaluate the difference between the relation of foster-child intelligence to true-parental factors and the relation to foster-parental factors.

In juxtaposing the results of different investigations, one is led to look for sources of differences and for possible sources of non-validity in the sets of data. It seems probable on the basis of present information that the proportion of illegitimate children determines in considerable part the general level of the sample, and this proportion has varied in different investigations. There is some question as to how justly we can assume that the true parents for whom information is available represent fairly the true parents for whom little or no data can be obtained. There is furthermore not much reason to put confidence in what data can be obtained on true fathers. For if it is a wise child who knows his own father, it is a brilliant illegitimate child who knows his 'true parents.' Two further considerations show the restricted application of findings with foster children: first, as is well known, in the population as a whole assortative mating produces a positive correlation in the intelligence of husbands and wives (52, 72), but there may be less of this effect among the parents of illegitimate children; and second, in no sense can foster children be said to represent a random sample of the population. Despite these difficulties, the foster-child studies represent the best evidence to date on the nature of the relation between socio-economic status and intelligence. If they do not possess the conclusiveness their authors may have hoped for, they are at least the best introduction to the field, and the reader is advised to study them and their commentaries (13, 49, 80, 90, 94, 96) for the many implications that cannot be presented here.

The main group of children studied by Freeman, Holzinger, and Mitchell (35) was called the "Home Group." The cases were secured from the chief society for the care and placement of dependent Protestant children in Illinois. Although the authors state that a very large proportion of the children placed by the Society are illegitimate, only 141 of the 401 cases in the Home Group are illegitimate. The illegitimate children were superior with respect to occupational background of true fathers, in cases where it was known, and in intelligence. The superiority in intelligence was present but not statistically significant when comparison was restricted to legitimate children who were committed to the care of the Society at an age comparable to the age when illegitimate children were committed. For the 401 children in the Home Group, the average age of commitment was 3 years, 6 months; the average age at placement in the foster home in which they were tested was 4 years, 2 months; and the average age when tested, 11 years, 0 months. The corresponding ages for legitimate and illegitimate children, considered separately, were, respectively, 4 years, 9 months, and 1 year, 4 months; 5 years, 8 months, and 1 year, 7 months; and 12 years, 2 months, and 9 years, 0 months. Thus the main group of the Chicago study is composed of two groups that are heterogeneous with respect to a number of factors crucial to the present problem. The foster fathers were definitely superior to Illinois males in occupational status, the true fathers just as definitely inferior. The inferiority of the true fathers was most marked with respect to the 177 legitimate fathers for whom information was available; the 78 illegitimate fathers for whom occupational information was available were more variable but not markedly inferior in occupational status to Illinois males.

Intelligence was measured by means of the Stanford-Binet and the International Group Mental Tests; only results obtained with the former will be reported here. Parents' intelligence was measured by means of the Otis Self-Administering Test, Higher Examination; vocabulary, by means of an unstandardized test. The home ratings were made by field-workers who gave equal weighting to six factors: material environment, evidences of culture, occupation of foster father, education of foster father, education of foster mother, social activity of foster parents.

The point on which the study has been criticized most severely is that there are a number of indications that selective placement, or the matching of child to foster family, took place.

Seventy-four children from the Home Group were tested just before placement in a foster home. Their average age at this time was 8 years, and the correlation between their intelligence and the home rating of the home in which they were placed was .34 ± .07. On retest, after an average interval of slightly over four years, this correlation was raised to .52 ± .06. The main results are summarized in Table III. The authors give no explanation of the selective factors that caused these correlations to be based on numbers ranging from 146 to 401.

In the same *Twenty-Seventh Yearbook* of this Society in which the study of Freeman, Holzinger, and Mitchell appeared is to be found the equally notable study of Burks (12).

Her subjects were 204 foster children, legally adopted, between 5 and 14 years of age at the time of being tested, placed in the foster homes before 12 months of age (average age, 3 months). Foster parents were white, non-Jewish, English-speaking, of American, British, or North European birth, as were the true parents, so far as known.

Control children were selected in a random fashion from school enrollments in such a way as to match the foster children for age and sex; control families were also matched for type of neighborhood and occupational class of father and were drawn from within the same 'natio-racial' limits. The mean Barr rating for the true fathers for whom there was sufficient information to make a judgment was 9.6, with a standard deviation of 3.1. Since the

TABLE III. — CORRELATIONS BETWEEN CHILD'S IQ AND FACTORS IN THE HOME *

Factor	*Foster Children*						*Control Children*			
	Freeman		*Burks*		*Leahy*		*Burks*		*Leahy*	
	N	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*
Home Rating	401	.48	206	.21	194	.19	104	.42	194	.53
Cultural Index			186	.25	194	.21	101	.44	194	.51
Father's Occupation	394	.37			194	.12			194	.45
Economic Status			181	.23	194	.15	99	.24	194	.37
Father's Education			173	.01	193	.16	102	.27	193	.48
Mother's Education			194	.17	192	.21	103	.27	194	.50
Mid-parent Education	?	.42			193	.20			194	.54
Father's Intelligence	180	.37	178	.07	178	.15	100	.45	175	.51
Mother's Intelligence	255	.28	204	.19	186	.20	105	.46	191	.51
Mid-parent Intelligence	169	.39	174	.20	177	.18	100	.52	173	.60
Father's Vocabulary	152	.27	181	.13	177	.22	101	.47	168	.47
Mother's Vocabulary	224	.37	202	.23	185	.20	104	.43	190	.49
Mid-parent Vocabulary	146	.36			174	.24			164	.56

* Adapted from Leahy (59, p. 287; and 58, p. 282) and Burks (13, p. 278). Probable errors are all between .03 and .05.

average IQ of children for whom Barr ratings of true fathers were available does not differ from that of foster children for whom Barr ratings could not be made, Burks infers the above value to be representative of the occupational background of the foster group. Terman (89) obtained an average of 7.9 and standard deviation of 3.5 for a random sample of adult males representing closely the same population as that from which the parents were drawn. For the foster fathers, the mean Barr rating at time of application for a foster child was 11.5 with a standard deviation of 2.6. Of the cases, 15 percent were known to be legitimate children, 72 percent known to be illegitimate, the rest doubtful. The Stanford-Binet test was administered to parents and children, homes were graded on an adaptation of the Whittier scale, and a culture index was computed, weighting equally parents' vocabulary (from Stanford-Binet results), education, interests, library, and artistic taste. As evidence against selective placement, Barr ratings of the true father showed zero correlations with Barr ratings of the foster father, Whittier ratings of the foster home, culture ratings of the foster home, and child's IQ at time of investigation.[1] The children were too young at time of placement to permit selection on the basis of their intelligence. The main results of this study are presented in Tables III and IV.

Leahy (58) studied a group of 194 adopted children similar in composition to those in Burks' research, making an even more rigorous attempt to avoid selective placement. The nature of the adoptive sample and carefully matched control group is described in Part I, Chapter XI of this Yearbook. Repetition of the following points is pertinent to the present discussion.

All the foster children were illegitimate; a preliminary study by Leahy showed that in Minnesota illegitimate children constitute a group more highly selected for intelligence than legitimate foster children. Leahy's attempt to get forty cases in each occupational category resulted in an excess of foster and control fathers in the higher occupational categories, particularly since she was unable to fill the quota for the two lowest classes. The variability of occupational status appears to be somewhat larger than in a random sample. The Stanford-Binet was given to all of the children; the parents took the Otis Self-Administering Test, Intermediate Form A, and the Stanford-Binet Vocabulary Test. Environmental status score and cultural index of home were based on homemade questionnaires. Father's occupation was rated on the Minnesota Scale for Occupational Classification. Although matched for occupational status and not significantly different in education, the foster homes were slightly superior to the control homes on the index of environmental status, which was composed of ratings of father's occupation, parents' education, economic status, degree of social participation, cultural materials, and child-training facilities at home.

Despite precautions against selective placement, a correlation of .18 appeared between occupation of true and adoptive fathers for eighty-nine cases, which is within chance limits. Correlations between true mother's occupation and education and factors in the foster home were smaller.

The main results are shown in Tables III and IV. For the adopted children, there are no statistically reliable differences in mean IQ between occupational classes, even for extreme groups. For the controls, the differences

[1] Does not the zero relation between child's IQ and Barr rating of the true father vitiate the reasoning on which it was concluded that the available Barr ratings represented the occupational distribution of all true fathers?

were statistically significant except for differentiation of first and last from adjacent groups. Leahy noted the smaller variability of foster-children's IQ as compared to control children and corrected the correlations for the difference in variability. Correlations reproduced in Table III are uncorrected, though the corrected ones differ by at most .05.

Table IV. — Child's IQ According to Father's Occupation *

	Foster			Control		
Father's Occupation	*N*	*Mean IQ*	*S.D.*	*N*	*Mean IQ*	*S.D.*
	Leahy Study					
I. Professional	43	113	11.8	40	119	12.6
II. Business Manager	38	112	10.9	42	118	15.6
III. Skilled Trades	44	111	14.2	43	107	14.3
V. Semiskilled	45	109	11.8	46	101	12.5
VI. and VII. Slightly Skilled and Day Labor	24	108	13.6	23	102	11.0
	Burks Study					
I. Professional, Higher Business	32	109	17.2	18	119	15.4
II. Semiprofessional	47	109	14.5	33	119	12.2
III. Lower Business	41	108	14.3	27	116	18.6
IV. Skilled Labor	43	105	16.7	18	106	12.4

* Adapted from Burks (13, p. 279).

Goodenough (Part I, Chapter XI, of this Yearbook) points to the possibility that selective factors may have curtailed the lower end of the distribution of intelligence of foster children. Another possibility is that the smaller variability of intelligence in foster children results from the fact that hereditary and environmental sources of variation are correlated in the control group but virtually uncorrelated in the foster group.[1]

Data and discussion of other foster-children studies will be found elsewhere in this Yearbook.

[1] The variance of a sum is equal to the sum of the variance of the constituents, in this case, heredity and environment, to a first approximation, plus twice the product of the correlation of the two constituents and the geometric mean of their variances. Where the two components have fixed variances, the variance of the sum will be greater or less depending on the correlation between them.

(2) Orphanage Children. A number of studies have appeared on the intelligence of children in orphanages.

Reymert and Frings (74) report a slight relation between intelligence of Mooseheart residents and the occupations of their deceased fathers, also a slight positive relation between intelligence and number of years in Mooseheart, which they consider a superior environment.

Lithauer and Klineberg (61) tested 120 children under the care of the Hebrew Orphan Asylum in New York City. About half the children had been put in the orphan asylum, half in foster homes. The authors attribute the average IQ gain of 5.9 points to the superiority of the new environment over the environmental origins of the children.

Recomputing Schott's (76) data, R. L. Thorndike (92) showed that a statistically significant rise in IQ of 4.0 points took place in the case of 74 children shifted from a very poor to a boarding-home environment. The statistical significance of the results does not, however, answer Schott's point that the rise is about the same as for retest studies not concerned with the effects of improved environment. As Goodenough (42) points out, Thorndike neglects also the unfavorable circumstances under which the initial tests were given. As the subjects were extremely heterogeneous with respect to age at time of first test, and as retesting was done only in cases selected for reasons not easily known, no generalizations can be made from his data

Skeels and Fillmore (83) have reported that the decline of IQ with age for a group of orphanage children (whom they consider to have a poor environment) is greater than the decline for unselected children. The unselected children were not, however, matched with the orphanage cases for original IQ level.

Jones and Carr-Saunders (51) used the Simplex, an English group test, to measure the intelligence of children living in eight schools, all selected because of the inadequacy of their home background. When their subjects, who were between the ages of 9 and 16, were classified into five groups according to father's occupation, a differentiation of intelligence similar to that found in the Northumberland study by Duff and Thomson appeared. There was some tendency for the differentiation between occupational classes and variability within occupational classes to decrease as length of residence increased. No statements of statistical significance were included. The report discussed insufficiently the effects of age at test, age at entrance, length of residence, and selective factors, and the interaction of these factors.

Lawrence (57) studied two groups of institutional children and two other groups, finding a marked decline in IQ with age for all groups, whether tested with the Simplex or Group Test 34 of the National Institute of Industrial Psychology. Her data, which are discussed in detail in Chapter XIII, showed sex differences in the degree of relation between father's occupation and intelligence. The unorthodox occupational classification, with professional people and all types of farmers thrown into a single category, makes the results difficult to compare with other studies reviewed here. Use of correlation ratios rather than product-moment correlations appears to compensate for the lowering of relations one would expect from the nature of the occupational classes.

(3) Twins. In the case of twins with identical biological heredity any differences in intelligence must be due to environmental sources, prenatal or postnatal.

An impressive battery of tests, including the Stanford-Binet, Otis Self-Administering, International Intelligence, American Council, and Stanford Achievement tests, was administered to 19 pairs of identical twins separated in infancy, in a study reported by Newman, Freeman, and Holzinger (68). Ratings of educational, social, and physical environmental differences were made by five judges, three of whom used only the case histories and two of whom also knew the twins. For the pooled judgments of the raters, all reliabilities were .90 or greater, but the smallest reliabilities were for the ratings of social environment. No discussion is given of the meaning of 'social environment.' They summarize the case histories by stating that for five pairs the differences in education and ability were large or fairly large and in the same direction; for one pair the difference in formal schooling was large without marked difference in ability appearing, possibly due to a compensating physical difference; one pair showed a marked difference in ability corresponding to marked differences in informal advantages, even though formal schooling was practically equivalent; the remaining twelve pairs showed small or negligible differences in education and ability. Because of the small number of cases, only correlations greater than .45 exceed the .05 level of significance. No significant relation appeared between physical and health differences and the intelligence differences. All the correlations between educational differences and intellectual ones were significant, the highest being the relations to Stanford Educational Age ($r = .9$) and Stanford-Binet IQ ($r = .8$). Social difference ratings correlated .5 with Binet IQ, Otis IQ, and the International Test, and .3 with the American Council Test and Stanford educational age. These results should be considered in connection with the fact that educational differences are probably more objectively rated than 'social' differences. A fuller discussion of data from research on twins is given in Chapter VIII of this Yearbook.

b. Interaction of Intelligence and Socio-Economic Status with Other Variables. The relations between intelligence and such factors as personality, size of family, race, and locality of residence are germane to the present discussion in so far as these factors are also related to socio-economic status. A complete survey of the interaction of the two major variables with other variables would have no limits. The studies included here are not claimed to be the most representative nor the most important. They have been selected partly by accident, partly because the interaction of the factors was explicitly discussed or the research was well designed to illustrate some point in the discussion. Attention will be focussed on the association of factors in the populations studied, without inference as to causal sequence. The reader should remember that except where they are explicitly ruled out, these factors may be associated with intelligence and socio-economic status in studies reported in other sections of this review.

(1) Rural-Urban Differences. The nature of rural-urban differences in intelligence may be essentially the same as the nature of differences between socio-economic groups within an urban or rural environment, or the two problems may be somewhat different. Occupational scales that include rural occupations rate them in the lower part of the scale. For occupations pursued both in rural and urban communities, it has been shown (26, and elsewhere) that, although the same sort of occupational differences in intelligence hold in rural localities as in urban ones, the rural members of the occupation are at least slightly inferior in intelligence to city-dwellers in the same occupational categories. Thus the problem of rural-urban differences does not seem to be solely one of occupational differences, though the differences in occupation and locale are almost inextricably associated.

Jones, Conrad, and Blanchard (53) compared rural-urban differences in intelligence with differences between children of average and superior socio-economic status within an urban environment, for each of the subtests of the Stanford-Binet in the year-levels 4 through 12. The rural group comprised an unselected sample of 351 5- to 14-year-old children from representative rural areas in northern New England. Average city children were represented by 921 cases from the 1917 Stanford standardization group. The superior urban selection consisted of 212 students in a country-day school in New York City, all between the ages of 6 and 14 years and of national-racial background comparable to that of the other two groups. It was estimated by the authors that the socio-economic status of the superior urban sample was above the status of 98 percent of the rural group and 90 percent of the urban group. The mean IQ of the rural group was 92; of the urban group, 101; and of the superior urban children, 117. The variabilities were similar for the three groups.

Among the evidences that the effects of environment are specific to the particular function tested are the facts that many of the tests are misplaced in relative difficulty for the rural group when a standardization based on urban children is used and that some tests show increasing, others diminishing, power to discriminate rural from urban groups. The tests showing rural-urban divergence were more frequent at higher age levels, and it was inferred that this is due to systematic changes in test content rather than to general tendencies in mental growth. For the individual test items, five showed marked rural-urban differences but slight urban-superior differences, and five showed reliable or nearly reliable urban-superior differences but negligible rural-urban differences. The authors suggest two explanations for these discrepancies, indicating that the explanation to be favored will depend upon the content of specific tests: (1) a curvilinear relation may exist between individual differences in the environment and in test performance, (2) certain test items may be relatively free from differential environmental effects, but in the measurement of intelligence may have a validity restricted to some part of the distribution of intelligence. The remaining 22 tests showed about equally reliable differentiation between rural-urban and urban-superior groups. For the 7 tests showing the most marked and reliable differences, three explanations are offered: (1) greater susceptibility of these tests to environmental influences, (2) greater hereditary differences in the functions measured by these tests, (3) greater validity of these tests in discriminating intellectual differences. The authors favor the first of these explanations. The third possibility, however, is not excluded by either of the first two, and depends on the purpose for which the validity is desired. On a form board involving rural content (the Mare and Foal Test), the rural children were definitely superior, whereas other performance tests showed various degrees of inferiority of the rural sample.

In the study of Iowa farm children made by Baldwin, Fillmore, and Hadley (4), it was found that the intellectual differentiation of rural from urban children and the differentiation of rural children in consolidated and one-room schools was absent in the preschool years and increased with age from 5 to about 14 years. Inferiority of rural to urban children was greater in language than in non-language subtests of the Stanford-Binet and was greater on the Stanford-Binet than on a series of performance tests. With a group of rural and city children matched both for mental and chronological age, performance on the Stanford-Binet subtests was studied. The results on specific subtests are somewhat similar to, but not identical with, the New England findings of Jones, Conrad, and Blanchard on the discriminative power of subtests.

Shimberg's (78) study, while concerned in the first instance with information rather than intelligence, is relevant by virtue of the fact that many items in intelligence tests are of essentially the same type as the items in her information test. This similarity was one of the motivating factors in the study, which was designed to test the validity of judging rural children on the basis of tests standardized on city children. For this purpose Shimberg collected information questions from both rural and urban school teachers. The most discriminative items submitted by urban teachers were selected and scaled on a group of over 400 city children, and the test, known as Information Test A, was standardized on over 6,000 city children. A similar procedure was followed in standardizing Test B on rural children. On Test A, the rural group was about a year retarded compared to the urban norms; on Test B, highly reliable differences in favor of rural groups appeared. The results are not so presented as to make it possible to compare the amount of urban superiority on Test A with the magnitude of rural superiority on Test B. But since the magnitudes depend so closely on the specific, and therefore accidental, content of the tests, the magnitudes of differences are not so important as the direction of differences shown. Although there was no measure of socio-economic status, different urban schools showed reliable differences in information scores.

(2) Race and Bilingualism. It has frequently been pointed out, probably first by Arlitt (2), that intellectual comparisons of 'races' or 'natio-races' do not always control the factor of socio-economic status. It is less usual to find objections against studies of socio-economic differences on the grounds that the race factor has not been controlled. We can only observe, however, that intellectual differences occur between races and classes, and that the races occur in different proportions in different classes. We cannot observe that any of the three variables is causal with respect to the others. It is not, moreover, easy to say just what would comprise a crucial experiment here. The statement that all factors other than socio-economic status and intelligence should be controlled is not as straightforward as it may sound. In many communities the selection of, say, second-generation Americans will not give at all a representative sampling of some occupational classes. There is therefore some reason for studying race concomitantly with intelligence and socio-economic status rather than attempting to hold race constant.

In one of the earliest studies of race differences, Arlitt (2) tested three groups of children taken from the primary grades of a single school district, native-born whites, Italians, and Negroes. Father's occupation was classified according to Taussig's categories and intelligence was measured by the Stanford-Binet. She noted that differences in intelligence for children of the same race but different social status were less than differences for different races within a single social class.

It is probable, however, that groups could be found of such variability as to reverse this difference; it should be further pointed out that race and social status acting together produce greater intellectual differences than either acting alone.

Third-grade Negro children of Washington, all of whom had been born in the District of Columbia, were tested by Long (62) on a battery including the Stanford-Binet, Kuhlmann-Anderson, Pintner-Paterson short scale, Form A of the Dearborn Group, paragraph meaning and word meaning from the New Stanford Reading tests, and reasoning and computation from the New Stanford Arithmetic tests. Subjects in Group I were 100 children from schools in a neighborhood that had been shown previously to be characterized by a rating of 13 on the Sims Socio-Economic Scale. Group II consisted of 100 children from schools in a neighborhood characterized by a Sims rating of 20. In other words, Group II was the superior socio-economic sample. In selecting Group II, principals and teachers were asked to recommend children from the better homes of the community, and it is possible that their recommendations were biased in favor of the more intelligent of the children from the privileged homes. When fathers' occupations were classified as professional, skilled, semiskilled, and unskilled, Group I fell slightly below the percentages for colored male adults in the city of Washington, whereas Group II was markedly above those norms. Group II was younger than Group I by an average of .4 years, but in terms of mental and educational ages all differences were in favor of Group II. The difference between mean mental ages of the two groups on the Stanford-Binet was .7 years, on the Kuhlmann-Anderson, .4 years; on the two performance tests, the mean differences were .9 years. For comparisons involving educational ages, the superiority of Group II ranged from .5 to .1 years. For 34 cases with Stanford-Binet IQ's greater than 120, all but two came from Group II, proportionately more came from higher occupational categories than for Group II as a whole, and the higher average IQ's belonged to children from higher occupational classes, even with this restricted range. The fact that performance tests showed sharper differentiation than verbal tests was borne out by analysis of the Pintner-Paterson and Kuhlmann-Anderson tests by subtests. All of the subtests of the performance scale showed differences in favor of Group II, and most of the differences were reliable. Half the subtests of the verbal scale showed differences in favor of each group. That performance tests should show greater differences than verbal tests is an atypical result, the meaning of which is not obvious. Possibly the answer is in terms of the absolute level of ability on which the various tests afford best differentiation; at any rate, the issue seems important enough to warrant further study.

A number of studies on Negro intelligence have been made by Klineberg (56). He has advocated the hypothesis that the superior showing of Northern Negroes on intelligence tests is an outcome of educational and socio-economic factors rather than of factors of selective migration. For the years 1914 to 1930, examination of the school records of Nashville and Birmingham showed that the pupils who migrated to the North were not superior to their classmates in school performance. Nine of Klineberg's graduate students tested Negroes who had been in New York for varying lengths of time, using the Stanford-Binet Test, National Intelligence Test, Intermediate Form of the Otis

Self-Administering Test, Pintner-Paterson Short Scale, Minnesota Paper Form Board, and Curtis Arithmetic Test. Different subjects were used for the different studies; in all, 3,081 10- and 12-year-old boys and girls in Harlem schools were subjects. Klineberg considers coming to New York to represent a rise in socio-economic status in many cases. The verbal tests in general showed a fairly consistent trend toward increase in score with increase in length of New York residence, but there was no evidence for relation between length of residence and performance on Curtis Arithmetic or performance tests. Two comments may be made on these studies: The only very clear results were obtained with the National Intelligence Test, the other studies suffering either from small number of cases or lack of consistent and reliable differences. If school facilities are as important as the author represents them to be, why was there no clearer rise in the results on the one achievement test used?

Six tests from the Pintner-Peterson scale were administered individually to the boys between the ages of 10 and 13 in Klineberg's (55) carefully conceived and widely quoted study on national and racial differences in Europe. Subjects were classified into Nordic, Alpine, and Mediterranean 'races' on the basis of region of residence and anthropological measurements. There were no consistent race differences, nor were the differences between boys from Paris, Hamburg, and Rome reliable. Rural-urban differences were marked, and consistently in favor of the cities.

Reviewing twenty-three reports concerning the comparative intelligence of Jews and non-Jews, Brill (9) found serious objections to each of the studies. He considered socio-economic status to have been ruled out in only three of the studies and language handicap in only one. As a general impression rather than a well-grounded conclusion, he stated that Jewish children in Great Britain and the United States are superior or at least equal to non-Jewish children of similar socio-economic status. If, however, Jewish families suffer a socio-economic handicap (which may or may not be the case), 'controlling' socio-economic status would result in increasing the superiority, if any, of the Jewish children.

Bilingualism is closely related to certain national-racial differences.

Arsenian (3) tested the students between the ages of 9 and 14 in two New York public schools, one of which was 90 percent Italian; the other, 90 percent Jewish. His sample included 1,152 native-born Italian subjects and 1,196 native-born Jews. Bilingualism was measured by means of the Hoffman Schedule, intelligence by means of the Pintner Non-language Test and Part I of the Spearman Visual Perception Test. Socio-economic status was measured by an adaptation of the Sims Score Card, which correlated .86 with the total Sims score and showed a reliability of .83 for 618 siblings. Father's profession was weighted heavily in the score. Bilingualism was found to continue in the second generation of American births for both groups, and the degree of bilingualism appeared to be symptomatic of adjustment to America. For both groups, longer residence in the United States resulted in higher socio-economic status, higher age-grade status, less bilingualism, and fewer children per family. Socio-economic status correlated negligibly with the measures of intelligence, slightly higher with age-grade status (.12 and .27). For whatever reason, all variables correlated higher in the Italian group than in the Jewish one, although the variabilities of the groups were in close agreement. For a sample of 67 foreign-born Italians, socio-economic status correlated .22 ± .08 and .26 ± .08 with the Pintner and Spearman tests, respectively. Correlations for a corresponding group of foreign-born Jews were negligibly negative. Within each of the native-born samples bilingualism correlated —.20 with socio-economic status.

Other studies (47, 48, 50, 71) seem to suggest that the low IQ of many bilinguals may be due to a combination of factors rather than language handicap alone.

(3) Birth Rate and Size of Family. It has frequently been reported that size of family is inversely correlated with social status. This finding has precipitated considerable agitation about 'dysgenic trends'; for the purposes of our review, it is unfortunate that little of the agitation has dealt with causal factors in the relation. For example, little is to be found on the subject of the implications for intellectual development of being brought up in a household predominantly adult as against a household composed predominantly of children. The present review will be limited to studies that present intelligence measures as well as socio-economic measures and vital indices.

An early study by Chapman and Wiggins (24) used the Chapman-Sims Socio-Economic Scale and the National Intelligence Test. All school children in Grades VI through VIII in a large manufacturing town along the New England coast were tested. The town has a considerable foreign population, especially Italians, Russians, and Poles. When consideration was limited to 'complete' families — that is, those in which the last birth was five years prior to the study — 632 cases remained in the sample. Socio-economic status correlated positively with IQ to the extent of .32, negatively with family size to the extent of —.27. The correlation between size of family and IQ was —.33. All the probable errors were between .02 and .03.

The factor of foreign birth was considered in a subsidiary study that involved matching 25 children of superior intelligence with a similar group of equal intelligence but with all parents foreign-born. The native-born group had an average socio-economic rating corresponding to the 80th percentile, whereas the foreign-born parents had an average social status rating in the 57th percentile. Thus the very superior foreign-born family did not rise much above the average socio-economic level of the community during the first generation.

In his study of intelligence and parental occupation, Collins (28) also obtained information on the number of children in the families. His subjects were 4,727 children from 3,089 families, enrolled in Grades I to VI of an Ohio industrial city with a population of 45,000. In Grades I through IV intelligence was measured with the Otis Primary Test; in Grades V and VI, with the Otis Advanced Group Test. For 218 children in special schools for the retarded, Stanford-Binet IQ's were available. IQ's from different tests were apparently considered as equivalent. Mean IQ's of his subjects, according to occupational class of father, are as follows: professional, 115; managerial, 113; clerical, 112; trade, 109; foreman, 106; skilled labor, 102; unskilled labor, 94; agricultural, 100; unknown, 96. The occupational distribution for the sample was approximately like that for the city as a whole. The mean number of children in the families from which his subjects were drawn was 3.4; in the professional group the mean number of children was 2.6; in the unskilled labor group, 4.1. Within each occupational group there was also a tendency for larger families to have lower IQ's.

In studying intelligence in the rural population of nine New England counties, Conrad and Jones (29) determined the size of the families and social status, as rated on a nine-point scale. They found no significant relation between the number of living offspring and such factors as parental intelligence (as ascertained with the Army Alpha test), parental education, and social status. Because child-bearing begins at an earlier age in the lower classes, there is a shorter interval between their generations. The advantage in reproduction so obtained is slight, and, the authors believe, probably counterbalanced by differential death rate in maturity.

In contrast with the above results from a relatively stable rural culture, in an urban population Burks and Jones (14) found correlations between IQ and size of sibship of −.22 and −.31. In Oakland the mean size of the sibship ranged from 2.0 for professional and executive families to 3.0 for unskilled and unemployed groups. That the relation was not linear is shown by the fact that two-child families have slightly higher IQ's than one-child families, and the fact that families in the skilled labor category are larger than those in the semiskilled class.

Cattell (19, 20), in his study of English populations, divided occupations into 14 groups and computed average values of intelligence, fertility, and earning power for each group. Correlations between these averages were .91 for intelligence and earning power; −.84 for intelligence and size of family; −.78 for fertility and earning power. Within the majority of occupations the size of the family was negatively related to intelligence. In the case of a non-verbal test, Cattell found that while the mean IQ of urban children was superior to that of rural children, the families were smaller in average size.

Within both samples, a decline in IQ of children was found with increasing size of family.

Residence in one of the 310 'health areas' of New York City was used to provide a measure of socio-economic status in the study made by Maller (65). The intelligence rating for each area was based on composite results obtained with the National Intelligence Test and Pintner Rapid Survey Test, given all fifth-grade students in New York City in one year (N = 100,153). Indices of birth rate, death rate, and infant mortality were derived from data in the city Department of Health; information about school retardation was obtained from school records. His correlations are based on areas, rather than on individuals. The number of areas is reduced to 275, apparently by combining some and leaving out others that were predominantly Negro in population. Intelligence correlated .70 with the index of school progress; .16 with school attendance; —.34 with birth rate; —.43 with death rate; —.51 with infant mortality; and —.57 with rate of juvenile delinquency. Contrary to expectation, the proportion of foreign-born in the area shows no marked relation to any of the other factors, though it is slightly positive with intelligence and slightly negative with delinquency rate. The proportion of Italians in an area correlated —.42 with the average intelligence; the proportion of Jews correlated .47 with intelligence. "The study indicates that the respective areas of the city have population units quite homogeneous in physical, mental, and social characteristics."

An able review of the relation of vital indices to social status and intelligence has been prepared by Lorimer and Osborne (63). They also present a detailed discussion of the interaction of these factors with racial, rural-urban, and regional differences.

(4) Personality and Behavioral Factors. Even though a general review of the relation of personality and behavioral factors to intelligence may reveal no consistent trends, the possible interaction of these factors with socio-economic status deserves consideration. It has been suggested that the relation of behavioral factors to intelligence may differ in the different social classes. If this is the case, groups heterogeneous with respect to social status will yield results in which the relationship is masked. It seems reasonable to suppose that personality factors are more closely related to test performance at preschool than at older levels, and that this fact possibly affects the findings concerning socio-economic and intellectual relations at the preschool level.

On the basis of clinical study of eleven matched pairs of children between the ages of 31 and 54 months, Gesell and Lord (38) conclude that the children from superior socio-economic classes are as much superior in personality factors as in intellectual development.

Springer (84), studying over 400 children each from a slum neighborhood and a good middle-class neighborhood, found large and significant differences in Barr ratings of fathers' occupations and in problem-tendency scores, judged from the Haggerty-Olson-Wickman Behavior Rating Schedules. There were almost no differences in average age (range from 6 to 12 years) or in intelligence as judged by the Goodenough Drawing-a-Man Test. Correlations between behavior ratings and intelligence within the social groups, computed for boys and girls separately, were all negative but not consistently reliable. Considering the heterogeneity with respect to age and homogeneity with respect to other factors, these correlations may be biased either in the direction of magnifying or underestimating the actual degree of relationship.

c. Analysis of Socio-Economic Status. The data of Table III may be interpreted as relevant to the problem of the relation of various aspects of socio-economic status to intelligence. The first four factors in that table refer to ratings that might be called ' socio-economic ' in a fairly strict sense. They are followed by the educational, intelligence, and vocabulary ratings, or scores, of the parents. Since the socio-economic, educational, intellectual, and vocabulary ratings of the parents are ordinarily closely related, the question arises as to which of these home factors is responsible for the relation of the group of factors to the child's intelligence. If the correlations are mainly due to biological heredity, we should expect the child's intelligence to be related more closely to the parents' intelligence than to the other factors. For the two control groups (' own ' children) the highest correlations in the column are those with mid-parent intelligence. (It is interesting to notice that the lowest correlations for the control groups are with measures of economic status — income in Burks' study and an economic index in Leahy's.) Should social stimulation be the chief factor in producing the relations, we should expect the child's intelligence to be related most closely to parents' vocabulary, since the vocabulary test is the only one that measures a direct medium of environmental transmission of intelligence. If selective placement is the chief source of the correlations between the intelligence of foster children and factors in the foster home, we should expect the highest correlations to be with socio-economic factors in the foster home, since they represent data most likely to be available to placement agencies. In the Freeman study, the highest correlation falls in the socio-economic category. In Leahy's study, the highest correlations are between child's intelligence and parents' vocabulary. Burks' foster children are inconsistently different from those of Leahy with respect to correlations with parents' vocabulary; they are consistently but very slightly higher in the relation between socio-economic factors and child's IQ. A possible explanation of this fact may be the more rigorous precautions against selective placement in Leahy's study. Before giving much attention to these correlation patterns, the reader is urged to consider the possibility that the variation in correlations within any one column is an outcome of accidents of sampling and errors of measurement. Real differences in the degree of relation of these factors to the child's intelligence probably exist, but the direction of those differences cannot be inferred with certainty from available data.

Comparable data at the preschool level have already been discussed in connection with the study of Bayley and Jones (6). It was noted that between two and six years the highest correlation was with mother's education, with other factors following in the order, father's education, father's occupation, total socio-economic rating, rating of home and neighborhood, and family income. Possibly the education of the parents more nearly measures their intelligence than any of the other factors.

A previous study (7) of high-school seniors showed an inconsistent relation between income of father and child's intelligence, thus confirming the low predictive value of income found by Burks (13), Leahy (59), and Bayley and Jones (6).

By various sampling methods, Van Alstyne (93) obtained a group of seventy-five children within three months of their third birthdays. All were white, born in English-speaking countries, living with their own parents in unbroken homes in an urban environment. The intelligence of the children was measured by the Kuhlmann-Binet Test; of their mothers, by the Thorndike Test of Word Knowledge. Socio-economic status was measured by the Chapin Scale for Rating Living-Room Equipment and by a questionnaire covering a large number of environmental items selected by experts as probably related to mental development. The mean Barr rating of the fathers' occupations was 10.4, with a standard deviation of 3.86. Both of these values are higher than those reported by Terman (89) for an urban population. The heterogeneity of the group may account in part for the magnitude of the intercorrelations shown in Table V.

Van Alstyne compares her correlations with corresponding ones in other studies, for mother's education, father's education, mother's vocabulary, home rating scale, number of books in child's library, and size of family. In every case, Van Alstyne obtained larger coefficients than the other investigators. Among the other correlations with child's mental age are mother's education, .60; father's education, .51; father's occupation, .50, and number of siblings, —.41. The probable errors of correlations reported above are between .03 and .06. Considering the foregoing factors and the numerous other more minute aspects of the environment, it was found that no environmental factor correlated significantly more with child's mental age than with mother's vocabulary, or vice versa.

> The factors which one might believe to have some effect on mental age, such as the number of constructive toys, books, hours read to, etc., have correlations which are the same as, or lower than, those which would not apparently have such an effect, such as cleanliness, height, possession of own bed, etc. (p. 48)

TABLE V. — CORRELATIONS BETWEEN ENVIRONMENTAL FACTORS AND INTELLIGENCE FOR THREE-YEAR-OLDS *

Factor	*Factor*			
	2	*3*	*4*	*5*
1. Child's Mental Age	.81	.54	.59	.61
2. Child's Vocabulary		.57	.67	.70
3. Mother's Vocabulary			.65	.70
4. Chapin Scale				.68
5. Environmental Composite				

* Adapted from Van Alstyne (93, p. 59).

IV. SUMMARY AND REORIENTATION

This brief review of the great amount of evidence on the relation between socio-economic status and intelligence can scarcely be condensed further without going beyond the limits justified by the data. Any summary statement necessarily resembles more nearly a series of most probable guesses than well-substantiated generalizations. With that qualification, and with the excuse that it is difficult to think in terms of discrete studies, the following summary is presented.

The correlation between the intellectual status of the father's occupation and the child's intelligence can be characterized approximately by a coefficient of .4. In terms of group averages, when the occupations are divided into about six categories, the highest or professional group is about one standard deviation above the mean of the population, whereas the lowest, or day-laborer class, is about half a standard deviation or less below the population mean.

This degree of relationship does not alter markedly with age from about 3 years through 18 years. Beyond 18 years it is difficult to obtain random samples or to know what constitutes a random sample. From birth until about 18 months the relation between developmental measures and socio-economic measures is slightly negative or zero.

The increase in relation of the two types of measures from 18 months until some time between 3 and 5 years accompanies the increasingly verbal content of the mental tests.

Results do not depend closely on the measure of general intelligence, whether group or individual, provided it is predominantly verbal in content. For performance tests, even such as purport to measure 'general intelligence,' the relation is found to diminish for most samples. The magnitude of this regression of occupational classes toward the mean when one shifts from verbal to performance tests is approximately the same as the magnitude of regression toward the mean of bright and dull groups, selected by criteria probably weighted in favor of verbal intelligence, when one tests them on performance tests.

There is little doubt that the degree of the relation varies with the measure of socio-economic status. The 'economic' part of the term appears to be gratuitous, both because most socio-economic scales are loosely related to income and because income predicts intelligence less adequately than almost any other environmental variable used. Whether because of unreliability or admixture of irrelevant factors, scales composed of economic, cultural, occupational, and educational factors are usually inferior in predictive value to straight measures of parental intelligence or education. Parental intelligence and education probably bear a somewhat closer relation to child's intelligence than does the intellectual status of the father's occupation.

The degree of relation characteristic of American school children holds approximately for English school children and has been confirmed with remarkable exactitude in a study of Russian school children. Sectional and rural-urban comparisons show shifts in the average level of intelligence but preservation of the socio-economic differentiation within each group. Average differences between sections of the country and between rural and urban groups may themselves be considered 'socio-economic' differences in a broad sense. The surety with which children from the professional class attain a mean IQ around 115 or 116 in practically every representative sample should be contrasted with the varying mean IQ's for the lowest occupational groups in different studies. In other words, the (tested) intellectual status of professional groups in different sections of the country varies less than the intellectual status of industrial and farm workers in the industries prevalent in the different regions.

The general trend of the relationship does not change when the range of either or both variables is restricted in an unbiased way, with one exception. Gifted children comprise a sample selected with limited range of intelligence; most samples of Negroes have less socio-economic variability than the population at large. Yet, even among superior Negro children, there is a definite positive relation between intelligence and father's occupational level. The single exception to date has been inmates of certain institutions for the feebleminded, among whom the relation of intelligence to occupational origin appears to be reversed, probably because of a combination of pathological cases and selective factors. In instances where the relation between the socio-economic status of the home and the child's intelligence is negligible, as in some of the foster-child and nursery-school studies, there is no reason to believe that the selection of parents from different social classes is uniformly rigorous. These groups probably constitute biased samples.

Stoke and Lehman (86), reviewing evidence on socio-economic relations to intelligence for a period mostly prior to that covered in the present review, draw about the same generalization as to the magnitude of the relation as that given at the beginning of this summary. They caution that the majority of gifted and superior children come from non-professional classes and from families of modest income, despite the fact that the relative frequency is greatest from the comparatively small professional class.

The data in the section on associations of intelligence and social status with other variables are illustrative rather than conclusive. They may serve, no less, to convince the reader that the nature of the relation between intelligence and social status cannot be resolved, once and for always, by the use of foster children, identical twins, or any single statistical technique. Undoubtedly the relation depends both on hereditary and environmental influences on development. For reasons stated at length, little confidence is placed on the quantitative estimates of the relative contributions of heredity and environment.

One commentator (79) has made the point that where variations in intelligence depend to a considerable measure on environment, educators have so far failed to develop many intelligences to their maximal capacity. When adequate education is provided for all, intellectual variations, Shuttleworth feels, will depend much more on hereditary differences, much less on environmental ones. The point is of course well taken, but let us permit an heretical objection against setting up any intelligence-test rating as representing the desideratum of development for all members of society.

'Conclusion' is "one word too often profaned" for this review to profane it. We are not at a concluding point. Research in the field has reached that stage of maturity at which one may profitably review and summarize, but not safely draw conclusions. Instead, for final statement, let us turn to the most promising lines for the extension of this research.

It has been stated frequently, as a moral of the type of finding reported above, that a mental test should be confined in its application to the group on which it was standardized and to closely comparable groups. Confining use to the standardization population does not, however, preclude the possibility of extraneous factors acting on the measurement systematically. Using a mental test only in those environments for which it was designed does not mean that social differences within that environment will not influence the score. Nor is it necessarily true that social differences greater than those in the original sample will influence the score. It is a matter for investigation within what limits social environment can be considered as constant in taking a given measurement. Further, a statement about the limits within which environment can be considered as constant should always be couched in terms of the purpose for which the test results are to be used. For example, the administration of the Stanford-Binet in English to a child who has never heard a word of English gives a valid score in the sense that it tells what he can do at that time and under those conditions. For the purpose of predicting his school performance in his own community, the test would have no validity. Again, the Stanford-Binet might give valid results for predicting the future school performances of two children raised in widely different environments, but predict poorly or not at all the relative intelligence of their offspring if raised in a common environment. From these considerations it is apparent that future research may fruitfully be directed toward stating within what limits and for what purposes social status can be considered constant in measuring intelligence with a given instrument.

Although the issue of general intelligence versus more specialized abilities cannot be considered settled, psychologists are becoming less willing to make broad statements about intelligence than they formerly were. The notion of intelligence as composed of several more or less independent 'factors,' while open to objection, affords at least as good a basis for studying the relation between intellectual traits and environmental variables as the older notion of 'general ability.' It is not clear in advance whether we would thereby throw light on the nature of factors or the nature of environmental influences; certainly we should be led to conclusions more specific and more conservative than those to which investigators of 'intelligence' as a whole come.

While no doubt the authors of the studies reported above have realized that 'socio-economic status' influences intelligence via some relatively narrow aspects or concomitants, they have mentioned that fact infrequently and have sought the specific medium of influence even more rarely. Again, it appears that results more concrete, more conservative, and more fruitful would result from investigation of specific concomitants of social status in relation to intellectual performances.

Three studies may be cited as illustrating these recommendations, though perhaps none of the authors would have repeated his procedure exactly if he had started more recently. Shimberg's (78) work represents one of the closest approximations to a specific definition of purpose. The study of Jones, Conrad, and Blanchard (53) exemplifies one analytic usage of 'intelligence.' The most comprehensive analysis of environment has been that of Van Alstyne (93). The reader who consults the originals of the articles cited in the survey of evidence will surely be impressed that these are three of the most effective studies in the field. Their effectiveness lies, not in the variety of groups studied or the quantity of data gathered, but rather in the clearness of conception and thoroughness of analysis.

In short, the majority of studies on the influence of socio-economic status on intelligence are criticizable for loose usage of the terms 'intelligence,' 'socio-economic status,' and 'influence,' and for failure to separate the influence of 'socio-economic status' from the influences of associated variables. It is recommended that future studies replace each of the terms by a more analytic usage and that they direct their research toward purposes with a more definite context than radical solution of the heredity-environment problem. Some workers have already demonstrated that these approaches can be pursued fruitfully.

References

(1) Anastasi, A. *Differential Psychology: Individual and Group Differences in Behavior.* (Macmillan, New York, 1937) 615 pp.

(2) Arlitt, A. H. "On the need for caution in establishing race norms." *Jour. Applied Psychol.,* 5: 1921, 179–183.

(3) Arsenian, S. *Bilingualism and Mental Development.* (Teachers College,

Columbia University, Contributions to Education, No. 712, New York, 1937) 164 pp.

(4) Baldwin, B. T., Fillmore, E. A., and Hadley, L. *Farm Children: An Investigation of Rural Child Life in Selected Areas of Iowa.* (D. Appleton-Century: New York, 1930) 337 pp.

(5) Ball, R. S. "The predictability of occupational level from intelligence." *Jour. Consult. Psychol.,* 2: 1938, 184–186.

(6) Bayley, N., and Jones, H. E. "Environmental correlates of mental and motor development: A cumulative study from infancy to six years." *Child Develop.,* 8: 1937, 329–341.

(7) Book, W. F. *The Intelligence of High-School Seniors: As Revealed by a Statewide Mental Survey of Indiana High Schools.* (Macmillan: New York, 1922) 371 pp.

(8) Bradway, K. P. "Paternal occupational intelligence and mental deficiency." *Jour. Applied Psychol.,* 19: 1935, 527–542.

(9) Brill, M. "Studies of Jewish and non-Jewish intelligence." *Jour. Educ. Psychol.,* 27: 1936, 331–352.

(10) Burdick, E. M. "A group test of home environment." *Arch. Psychol.,* 16: 1928, 115 pp.

(11) Burks, B. S. "Nature and Nurture: Their Influence upon Achievement." *The Twenty-Seventh Yearbook, Part II,* of this Society, 1928. (pp. 248–353)

(12) Burks, B. S. "Nature and Nurture: Their Influence upon Intelligence." *The Twenty-Seventh Yearbook, Part I,* of this Society, 1928. (pp. 219–316)

(13) Burks, B. S. "On the relative contributions of nature and nurture to average group differences in intelligence." *Proc. Nat. Acad. Sci.,* 24: 1938, 276–282.

(14) Burks, B. S., and Jones, H. E. "A study of differential fertility in two California cities." *Human Biol.,* 7: 1935, 539–554.

(15) Burks, B. S., and Kelley, T. L. "Nature and Nurture: Their Influence upon Intelligence." *The Twenty-Seventh Yearbook, Part I,* of this Society, 1928. (pp. 9–38)

(16) Byrns, R., and Henmon, V. A. C. "Parental occupation and mental ability." *Jour. Educ. Psychol.,* 27: 1936, 284–291.

(17) Canady, H. G. "The intelligence of Negro college students and parental occupation." *Amer. Jour. Sociol.,* 42: 1936, 388–389.

(18) Cattell, R. B. "Occupational norms of intelligence, and the standardization of an adult intelligence test." *Brit. Jour. Psychol.,* 25: 1934–1935, 1–28.

(19) Cattell, R. B. *The Fight for Our National Intelligence.* (P. S. King: London, 1937) 164 pp.

(20) Cattell, R. B. "Some further relations between intelligence, fertility, and socio-economic factors." *Eugenics Rev.,* 29: 1937, 171–179.

(21) Chapin, F. S. "A quantitative scale for rating the home and social environment of middle-class families in an urban community: A first approximation to the measurement of socio-economic status." *Jour. Educ. Psychol.,* 19: 1928, 99–111.

(22) Chapin, F. S. *The Measurement of Social Status by the Use of the Social Status Scale.* (University of Minnesota Press: Minneapolis, 1933) 16 pp.

(23) Chapman, J. C., and Sims, V. M. "The quantitative measurement of certain aspects of socio-economic status." *Jour. Educ. Psychol,* 16: 1925, 380–390.

(24) Chapman, J. C., and Wiggins, D. M. "Relation of family size to intelligence of offspring and socio-economic status of family." *Ped. Sem.,* 32: 1925, 414–421.

(25) Chauncey, M. R. "The relation of the home factor to achievement and intelligence-test scores." *Jour. Educ. Res.,* 20: 1929, 88–90.

(26) Clark, C. D., and Gist, N. P. "Intelligence as a factor in occupational choice." *Amer. Sociol. Rev.,* 3: 1938, 683–694.

(27) Coffey, H. S., and Wellman, B. L. "The rôle of cultural status in intelligence changes of preschool children." *Jour. Exper. Educ.,* 5: 1936–1937, 191–202.

(28) Collins, J. E. "The intelligence of school children and paternal occupation." *Jour. Educ. Res.,* 17: 1928, 157–169.

(29) Conrad, H. S., and Jones, H. E. "A field study of the differential birth rate." *Jour. Amer. Statistical Assoc.,* 27: 1932, 153–159.

(30) Cuff, N. B. "Relationship of socio-economic status to intelligence and achievement." *Peabody Jour. Educ.,* 11: 1933, 106–110.

(31) Davidson, P. E., and Anderson, H. D. "Are Edwards' socio-economic levels economic?" *School and Soc.,* 48: 1938, 153–156.

(32) Dubnoff, B. "A comparative study of mental development in infancy." *Ped. Sem., and Jour. Genet. Psychol.,* 53: 1938, 67–73.

(33) Duff, J. F. "Children of high intelligence: A follow-up inquiry." *Brit. Jour. Psychol.,* 19: 1929, 413–439.

(34) Duff, J. F., and Thomson, G. H. "The social and geographical distribution of intelligence in Northumberland." *Brit. Jour. Psychol.,* 14: 1923, 192–198.

(35) Freeman, F. N., Holzinger, K. J., and Mitchell, B. C. "Nature and Nurture: Their Influence upon Intelligence." *The Twenty-Seventh Yearbook, Part I,* of this Society, 1928. (pp. 103–218)

(36) Furfey, P. H. "The relation between socio-economic status and intelligence of young infants as measured by the Linfert-Hierholzer scale." *Ped. Sem. and Jour. Genet. Psychol.,* 35: 1928, 478–480.

(37) Furfey, P. H., and Muehlenbein, J. "The validity of infant intelligence tests." *Ped. Sem. and Jour. Genet. Psychol.,* 40: 1932, 219–233.

(38) Gesell, A., and Lord, E. E. "A psychological comparison of nursery-school children from homes of low and high economic status." *Ped. Sem. and Jour. Genet. Psychol.,* 34: 1927, 339–356.

(39) Glass, L. C. "The relation of the 'intelligence' of college students to the occupation of their parents." *Eugen. News,* 21: 1936, 1–2.

(40) Goodenough, F. L. "The relation of the intelligence of preschool children to the education of their parents." *School and Soc.,* 26: 1927, 54–56.

(41) Goodenough, F. L. "The relation of the intelligence of preschool children to the occupation of their fathers." *Amer. Jour. Psychol.,* 40: 1928, 284–294.

(42) Goodenough, F. L. "A comment on Robert L. Thorndike's 'Note on IQ changes in foster home children,' by Emmett L. Schott." *Jour. Applied Psychol.,* 23: 1939, 308–309.

(43) Goodenough, F. L., and Anderson, J. E. *Experimental Child Study.* (D. Appleton-Century: New York, 1931) 546 pp.

(44) HAUGHT, B. F. "The relation of intelligence of college Freshmen to parental occupation." *Psychol. Monog.*, 50: 1938, No. 5, 203–210.

(45) HERTZMAN, M. "The effects of the relative difficulty of mental tests on patterns of mental organization." *Arch. Psychol.*, 28: 1935–1936, 69 pp.

(46) HILDRETH, G. "Occupational status and intelligence." *Personnel Jour.*, 13: 1934, 153–157.

(47) HILL, H. S. "Correlation between IQ's of bilinguals at different ages on different intelligence tests." *School and Soc.*, 44: 1936, 89–90.

(48) HILL, H. S. "The effect of bilingualism on the measured intelligence of elementary-school children of Italian parentage." *Jour. Exper. Educ.*, 5: 1936–1937, 75–78.

(49) HOLZINGER, K. J. "The statistical evaluation of nature and nurture." *Jour. Amer. Statistical Assoc.*, 30: 1935, 274–280.

(50) JAMIESON, E., and SANDIFORD, P. "The mental capacity of the southern Ontario Indians." *Jour. Educ. Psychol.*, 19: 1928, 536–551.

(51) JONES, D. C., and CARR-SAUNDERS, A. M. "The relation between intelligence and social status among orphan children." *Brit. Jour. Psychol.* (Gen. Sect.), 17: 1926–1927, 343–364.

(52) JONES, H. E. "Homogamy in intellectual abilities." *Amer. Jour. Sociol.*, 35: 1929, 369–382.

(53) JONES, H. E., CONRAD, H. S., and BLANCHARD, M. B. "Environmental handicap in mental-test performance." *Univ. Calif. Publ. Psychol.*, 5, No. 3: 1932, 63–99.

(54) JORDAN, A. M. "Paternal occupations and children's intelligence scores." *Jour. Applied Psychol.*, 17: 1933, 103–119.

(55) KLINEBERG, O. "A study of psychological differences between 'racial' and national groups in Europe." *Arch. Psychol.*, 20: 1931, 58 pp.

(56) KLINEBERG, O. *Negro Intelligence and Selective Migration.* (Columbia University Press: New York, 1935) 66 pp.

(57) LAWRENCE, E. M. "An investigation into the relationship between intelligence and environment." *Brit. Jour. Psychol., Monog. Suppl.*, 16: 1931, 80 pp.

(58) LEAHY, A. M. "Nature-nurture and intelligence." *Genet. Psychol. Monog.*, 17: 1935, 235–308.

(59) LEAHY, A. M. "A study of adopted children as a method of investigating nature-nurture." *Jour. Amer. Statistical Assoc.*, 30: 1935, 281–287.

(60) LEAHY, A. M. *Measurement of Urban Home Environment.* (University of Minnesota Press: Minneapolis, 1936) 70 pp.

(61) LITHAUER, D. B., and KLINEBERG, O. "A study of the variation in IQ of a group of dependent children in institutions and foster homes." *Ped. Sem. and Jour. Genet. Psychol.*, 42: 1933, 236–242.

(62) LONG, H. H. "Test results of third-grade Negro children selected on the basis of socio-economic status, I." *Jour. Negro Educ.*, 4: 1935, 192–212.

(63) LORIMER, F., and OSBORN, F. *Dynamics of Population; Social and Biological Significance of Changing Birth Rates in the United States.* (Macmillan: New York, 1934) 461 pp.

(64) MACMURRAY, D. "A comparison of the intelligence of gifted children and of dull-normal children measured by the Pintner-Paterson scale, as against the Stanford-Binet scale." *Jour. Psychol.*, 4: 1937, 273–280.

(65) MALLER, J. B. "Vital indices and their relation to psychological and social factors: A study of 310 health areas in New York City with reference to birth rate, death rate, juvenile delinquency, school progress, and intelligence." *Human Biol.*, 5: 1933, 94–121.

(66) MURPHY, G., MURPHY, L. B., and NEWCOMB, T. M. *Experimental Social Psychology; An Interpretation of Research upon the Socialization of the Individual.* (Harper and Brothers: New York and London, 1937) 1121 pp.

(67) NEFF, W. S. "Socio-economic status and intelligence: A critical survey." *Psychol. Bull.*, 35: 1938, 727–757.

(68) NEWMAN, H. H., FREEMAN, F. N., and HOLZINGER, K. J. *Twins: A Study of Heredity and Environment.* (University of Chicago Press: Chicago, Ill., 1937) 369 pp.

(69) PATERSON, D. G. *Minnesota Mechanical Ability Tests; the Report of a Research Investigation Subsidized by the Committee on Human Migrations of the National Research Council and Conducted in the Department of Psychology of the University of Minnesota.* (The University of Minnesota Press: Minneapolis, Minn., 1930) 586 pp.

(70) PATERSON, D. G., and RUNDQUIST, E. A. "The occupational background of feeblemindedness." *Amer. Jour. Psychol.*, 45: 1933, 118–124.

(71) PINTNER, R., and ARSENIAN, S. "The relation of bilingualism to verbal intelligence and school adjustment." *Jour. Educ. Res.*, 31: 1937–1938, 255–263.

(72) POPENOE, P. "Assortative mating for occupational level." *Jour. Soc. Psychol.*, 8: 1937, 270–274.

(73) PROCTOR, W. M. "Intelligence and length of schooling in relation to occupational levels." *School and Soc.*, 42: 1935, 783–786.

(74) REYMERT, M. L., and FRINGS, J. "Children's intelligence in relation to occupation of father." *Amer. Jour. Sociol.*, 41: 1935, 351–354.

(75) SANGREN, P. V. "Comparative validity of primary intelligence tests." *Jour. Applied Psychol.*, 13: 1929, 394–412.

(76) SCHOTT, E. L. "IQ changes in foster-home children." *Jour. Applied Psychol.*, 21: 1937, 107–112.

(77) SCHWESINGER, G. C. *Heredity and Environment: Studies in the Genesis of Psychological Characteristics.* (Macmillan: New York, 1933) 484 pp.

(78) SHIMBERG, M. E. "An investigation into the validity of norms with special reference to urban and rural groups." *Arch. Psychol.*, 16: 1928–1929, 84 pp.

(79) SHUTTLEWORTH, F. K. "The nature versus nurture problem. Part I. Definition of the problem." *Jour. Educ. Psychol.*, 26: 1935, 561–578.

(80) SHUTTLEWORTH, F. K. "The nature versus nurture problem. Part II. The contributions of nature and nurture to individual differences in intelligence." *Jour. Educ. Psychol.*, 26: 1935, 655–681.

(81) SIMS, V. M. *The Measurement of Socio-Economic Status: In Coöperation with the Character Education Inquiry, Teachers College, Columbia University.* (Public School Publishing Company: Bloomington, Ill., 1928) 33 pp.

(82) SIRKIN, M. "The relation between intelligence, age, and home environment of elementary-school pupils." *School and Soc.*, 30: 1929, 304–308.

(83) SKEELS, H. M., and FILLMORE, E. A. "Mental development of children from

underprivileged homes." *Ped. Sem. and Jour. Genet. Psychol.*, 50: 1937, 427–439.

(84) Springer, N. N. "The influence of general social status on school children's behavior." *Jour. Educ. Res.*, 32: 1938–1939, 583–591.

(85) Stoke, S. M. *Occupational Groups and Child Development: A Study of the Mental and Physical Growth of Children in Relation to Occupational Grouping of Parents.* (Harvard University Press: Cambridge, Mass., 1927) 92 pp. (Harvard Monographs in Education, No. 8.)

(86) Stoke, S. M., and Lehman, H. C. "Intelligence-test scores of social and occupational groups." *School and Soc.*, 31: 1930, 372–377.

(87) Stroud, J. B. "A study of the relation of intelligence-test score of public-school children to the economic status of their parents." *Ped. Sem. and Jour. Genet. Psychol.*, 35: 1928, 105–111.

(88) Taussig, F. W. *Principles of Economics.* (Macmillan: New York, 2d ed., 1915) Vol. II, 576 pp.

(89) Terman, L. M., Baldwin, B. T., Bronson, E., and others. *Genetic Studies of Genius: Vol. I. Mental and Physical Traits of a Thousand Gifted Children.* (Stanford University Press: Stanford, Calif., 1925) 648 pp.

(90) Terman, L. M., Freeman, F. N., Gates, A. I., and others. "Nature and nurture: Their influence upon intelligence and upon achievement." (Selected papers read at the Boston meeting of this Society.) *Jour. Educ. Psychol.*, 19: 1928, 361–409.

(91) Terman, L. M., and Merrill, M. A. *Measuring Intelligence: A Guide to the Administration of the New Revised Stanford-Binet Tests of Intelligence.* (Houghton Mifflin: Boston, Mass., 1937) 461 pp.

(92) Thorndike, R. L. "Note on 'IQ changes in foster-home children' by Emmett L. Schott." *Jour. Applied Psychol.*, 22: 1938, 662.

(93) Van Alstyne, D. *The Environment of Three-Year-Old Children: Factors Related to Intelligence and Vocabulary Tests.* (Teachers College, Columbia University, Contributions to Education, No. 366, New York, 1929) 108 pp.

(94) Wallis, W. D. "Observations on Leahy's 'Nature-Nurture and Intelligence.'" *Ped. Sem. and Jour. Genet. Psychol.*, 49: 1936, 315–324.

(95) Williams, J. H. "The Whittier scale for grading home conditions." *Jour. Delinq.*, 1: 1916, 273–286.

(96) Wright, S. "Statistical methods in biology." *Jour. Amer. Statistical Assoc.* (Suppl.), 26: 1931, 155–163.

(97) Yerkes, R. M. [editor]. "Psychological examining in the United States Army." [In] *Memoirs, National Academy of Sciences,* XV, Part III, Chapter 15. (Government Printing Office: Washington, D. C., 1921) pp. 819–837. See also simpler presentation of this material in Yoakum, C. S., and Yerkes, R. M. [editors]. *Army Mental Tests.* (Holt: New York, 1920) 303 pp.— *Editor.*

CHAPTER VI

SEX DIFFERENCES IN INTELLIGENCE-TEST SCORES [1]

G. M. Kuznets
Formerly Instructor in Psychology
and
Olga McNemar
Stanford University
Palo Alto, California

Within the last forty years a considerable amount of quantitative material has accumulated bearing on the problem of sex differences in intelligence. Much of the information may be found summarized in a number of reviews, the more recent and extensive of which are Goodenough (4), Lincoln (7), McNemar and Terman (8), Miles (10), and Wellman (18). The present chapter will survey only a limited portion of the available material, since it is restricted to a consideration of results obtained on total scores of standardized intelligence tests. The review is further restricted, as it will ignore evidence obtained from (*a*) admittedly selected samples, such as high-school and, more especially college, groups; (*b*) groups exhibiting wide variation in age and in which classification by age is not possible; (*c*) groups of inadequate size (few studies numbering less than 100 of each sex will be reviewed); (*d*) studies in which tests of significance cannot be applied. Finally, since it is doubtful whether differences in skewness and kurtosis can be given a meaningful interpretation, comparisons will be confined to measures of central tendency and dispersion.

I. Evidence from Verbal or Largely Verbal Tests

Most of the available information concerning sex differences in intelligence comes from studies based on samples of school children. Table I summarizes the results abstracted from ten studies of this kind. In evaluating these data, it should be kept in mind that school samplings are restricted in range of ability, that selective factors definitely operate in the upper grades, that a wide grade range is necessary

[1] This is a partial report of a larger study of psychological sex differences made possible by a grant to Stanford University on the recommendation of the Committee for Research on Problems of Sex, of the National Research Council.

TABLE I. — SEX DIFFERENCES IN INTELLIGENCE SCORES SHOWN IN TEN STUDIES OF SCHOOL CHILDREN

All critical ratios are in terms of standard errors. A positive ratio indicates a difference in favor of boys. Standard errors and critical ratios (except for Marshall) computed by the present writers. For critical ratios of the standard deviation, see McNemar and Terman (Ref. 8).

Investigator and Date	*Group*	*Test*		*Age, Years* 5	6	7	8	9	10	11	12	13	14	15	16	17
Terman (16) 1917	First Stanford Revision Standardization Group	First Stanford Revision of the Binet Scale	Boys	87		100		92		74		106				
			Girls	87		91		108		88		74				
			CR_M	−2.5		− .2		.0		−1.6		− .5				
			$CR\sigma$	− .3		− .2		−2.5		.3		− .5				
Pressey (12) 1918	Indiana School Children Grade III through High School	Pressey Intelligence Test	Boys				57	132	176	179	182	174	138	102		
			Girls				92	153	177	165	180	174	163	139		
			CR_M				−1.8	−1.1	−2.3	−1.7	−1.6	−2.6	−2.9	−2.8		
			$CR\sigma$				1.8	1.6	2.3	2.2	1.9	2.8	4.5	2.9		
Whitmire (20) 1920	Unselected School Children in Vallejo, California	National Intelligence Test	Boys				92	95	95	98	89	89				
			Girls				79	102	77	88	98	67				
			CR_M				−1.7	−1.6	−1.4	−2.0	−1.0	−2.3				
			$CR\sigma$				.1	.4	1.4	− .5	− .7	1.7				
Thorndike (17) 1926	Grades VIII, IX, and X in 3 Cities: City 1, 1922	IER	Boys									180	615	775	728	385
			Girls									198	716	1054	865	436
			CR_M									2.0	1.9	2.2	1.5	4.5
			$CR\sigma$									.5	.7	1.2	− .4	.2
Thorndike (17) 1926	Grades VIII, IX, and X in 3 Cities: City 2, 1922	IER	Boys										220	377	373	153
			Girls										274	494	405	189
			CR_M										3.2	1.9	3.1	3.4
			$CR\sigma$										.2	1.6	1.2	.3

Thorndike (17) 1926	Grades VIII, IX, and X in 3 Cities: City 2, 1924	IER	Boys									252	555	439	170
			Girls									279	637	490	155
			CR_M									.0	.5	— .7	1.1
			$CR\sigma$									1.0	1.3	2.0	.3
Whipple (19) 1927	11-year-olds from cities in Mich., Mo., and Ind.	National Intelligence Test	Boys							1071					
			Girls							1127					
			CR_M							—4.31					
			$CR\sigma$							less than 1					
Commins (1) 1928	Children of small city with enroll-ment about 3,000	McCall Multi-mental	Boys					92	105	94	102	97	70		
			Girls					98	119	98	101	84	70		
			CR_M					1.6	1.5	1.0	2.3	2.6	1.7		
			$CR\sigma$					1.0	1.6	—2.1	.8	1.6	.0		
Hales (5) 1932	A sampling of schools in Sydney, Australia	Advanced Test of General Intelligence	Boys								123	314	349	268	128
			Girls								122	310	284	232	170
			CR_M								—1.9	—1.7	—1.8	4.4	4.3
			$CR\sigma$								—2.1	— .8	2.4	— .1	— .3
Heilman (6) 1933	10-year-old children from 48 Denver schools	Stanford-Binet	Boys						464						
			Girls						482						
			CR_M						1.7						
			$CR\sigma$						1.3						
Marshall (9) 1934	State school children in Western Australia	Spearman's measure of intelligence	Boys				201	370	552	593	477	556	463	257	
			Girls				200	383	545	549	499	491	370	122	
			CR				.6	.0	— .1	.7	1.9	1.7	1.4	1.8	
McNemar and Terman (8) 1936	Kg. and Grade I, Berkeley, California (From Dickson)	Stanford-Binet	Boys	543	795	177									
			Girls	504	779	110									
			CR_M	7.6	—1.2	1.5									
			$CR\sigma$	1.7	— .7	.7									

in order to permit even the central ages therein to be representative age samplings, and that exclusions due to private-school enrollments introduce unknown selective factors.

Data presented in Table I seem to justify the opinion prevalent in psychological literature concerning the relative equality of the sexes in average mental status. For the available 56 comparisons of means or medians, 19 yielded critical ratios of 2 or greater, and only 8 of 3 or greater. Consistent trends, however, are observable on certain tests. Thus for the Pressey test, while only three of the differences may be considered reliable, all the mean differences favor girls. A similar superiority of girls is found on the National Intelligence Test, as indicated by the studies of Whitmire (20) and Whipple (19). On the other hand, a consistent superiority of boys is evident for two of the three groups in Thorndike's study (17) based on somewhat older children and also on the McCall Multimental Scale reported by Commins (1).

One can only conjecture concerning the meaning of the observed consistencies. Selective factors probably account for the differences shown in Thorndike's study — an explanation that Thorndike himself favors (17). It is not likely, however, that such selective factors are responsible for the superiority of the girls on the Pressey Tests and National Intelligence Tests. A more plausible conjecture is that this superiority is a function of the specific content of these tests. It is possible, for example, that primarily verbal components are given undue weight in the determination of the total, a procedure that would favor the more linguistic sex. While specific evidence is available on this point, it need not be cited here. In view of the relatively small magnitude of the differences, it seems possible that changes in the content of the test that would not impair its effectiveness as a whole might erase the difference or even change its direction.

Differences in dispersion fail to reveal significant sex trends. Of the 48 available comparisons, only 10 yield critical ratios of 2 or greater, and only one of 3 or greater. Although most of the differences are in the direction of larger dispersion for boys (34 out of 48), only the Pressey study yields consistent results.

The studies thus far considered are based upon groups that do not include extreme variants. Two recent studies may be cited in which the range of ability sampled is practically complete. The Scottish Mental Survey is remarkable for its extensiveness and the extreme care with which the data have been collected. The 1933 report (15) summarizes the results obtained on a verbal group test of an omnibus type and the Stanford-Binet. The group test was administered to the total age group of Scottish children 10½ to 11½ years old (except the blind and the deaf), comprising altogether over 87,000 children. While the difference between the means was not reliable, boys were found to be significantly more variable than girls (the difference between standard deviations was 18 times its probable error). The Stanford-Binet was given to a subsample of 500 children of each sex. After applying a correction to free the sample from bias due to over-representation of superior children, the report concluded that the sex difference between the mean IQ's was not significant, but that there probably is "a real difference in scatter of IQ's between boys and girls, a difference which seems from all the evidence to be something like one point in IQ in standard deviation." (The boys have the greater variability.)

The latter conclusion is not substantiated by more recent evidence obtained upon a completely unselected sample.[1] With the exception of one boy, every child in Scotland born on one of four specified days in 1926 was given a Stanford-Binet and a performance battery. At the end of testing, the age range of the group (consisting of 444 boys and 430 girls) was approximately 9 to 11 years, with an average age of 10 years, 5 months. It will be seen with reference to the following tabulation, which presents the results for Binet testing, that differences between the sexes in both central tendency and variability are not statistically reliable.

	Mean	*S.D.*	CR_M	CR_σ
Boys	100.51	15.88	.86	.85
Girls	99.70	15.25		

In the study by Roberts and others (13) the sample was defined as consisting of children born between specified dates in 1921 and 1924 whose homes on a specified day in 1934 fell within the boundaries of the city of Bath (England). The great majority of the ages of the 1,336 boys and 1,217 girls tested on the Advanced Otis ranged from 9 years, 10 months, to 12 years, 10 months, with a scattering down to 9 years, 5 months, and up to 13 years, 5 months. The defective children were not actually tested, but their number was ascertained and they

[1] The writers are greatly indebted to the Scottish Council for Research in Education for permission to quote from a confidential report on the preliminary results of a truly remarkable survey.

were assigned an arbitrary score on the test. The pertinent information is cited in the tabulation below.[1] Again the difference between means is shown to be statistically unreliable; the difference between the standard deviations, however, especially for fixed age, tends toward significance.

	Sex	*Mean*	*S.D.*	CR_M	$CR\sigma$
Age varying	Boys	85.07	39.41	.36	2.08
	Girls	84.53	37.18		
Age constant *	Boys	84.37	36.28		2.67
	Girls	84.09	33.66		

* Only variability accounted for by linear regression is eliminated. Means at fixed age are at 11 years, 4 months.

A discussion of sex differences in intelligence is hardly complete without some mention of possible differences not only in mental growth, but also in decline of ability. Data on sex differences, as measured by the Army Alpha, are presented by Conrad, Jones, and Hsiao (2) for groups ranging from early adolescence to old age. Although slight selective factors were undoubtedly operating at certain age levels, the group as a whole may possibly be regarded as a fairly representative sample of a rural New England community. For the four age groupings, 10 to 15, 16 to 21, 22 to 39, and 40 to 59, the differences between the means favor the females. For two of these groupings the differences may be considered significant (critical ratio 5.3 for the 10- to 15-year group, and 2.6 for the 40- to 59-year group). However, to quote from the authors, " In the range of ages from ten to sixty, the impressive fact is not the degree of sex differences, but rather the similarity of the developmental curves for the two sexes."

II. Evidence from Non-Verbal Tests

Relatively little acceptable evidence is available on non-verbal scales. In the Scottish Mental Survey previously described, the children were given a performance battery consisting of the following eight tests: Seguin Formboard, Manikin, Stutsman Four-piece Picture Test, Red Riding Hood (star insertion), Healy Picture Completion, Knox Cube Imitation, Cube Construction, and Kohs Block Design. The correlation between the performance battery and the Stanford-Binet mental age, with age held constant, was .67 for boys and .71 for girls. The results obtained on the performance battery are tabulated here-

[1] Standard deviations and critical ratios calculated by the present writers.

with. Significant differences in both means and standard deviations are found, in each case the boys' statistic being the larger of the two.

	Sex	*Mean*	*S.D.*	CR_M	$CR\sigma$
Age varying	Boys	56.32	15.83	3.74	2.77
	Girls	52.55	13.93		
Age constant *	Boys		15.05		2.16
	Girls		13.56		

* Only variability accounted for by linear regression is eliminated.

In Table II are reported data on Goodenough's Draw-a-Man Test, Dearborn Examination, and the Pintner Non-Language Test based on acceptably large age-sex groups. Here again impressive sex differences fail to emerge. The small, but consistent, differences in favor of girls on Goodenough's test are probably attributable to the method of standardization; the individual items were weighted in proportion to their correlation with school progress. Since available information indicates a differential in school progress in favor of girls, one might expect girls to excel on items thus weighted. Further evidence on five non-verbal tests is available from Schiller's study (14) based on a sample of 189 boys and 206 girls from the third and fourth grades of a New York public school. Girls were found to be superior on the Goodenough test. The Otis Primary and the Army Beta tests yielded significant mean differences in favor of boys (the critical ratios were 3.4 and 4.6, respectively). The difference between means on the Pintner-Paterson Scale, the International Test, and the differences between standard deviations on all five tests were not reliable.

III. Conclusion

A survey of the results abstracted from acceptable studies has yielded largely negative conclusions. When large unselected groups are used, when age is taken into account, when possibilities of bias in test content are allowed for, startling differences between the sexes either in average tendency or in variation fail to emerge. While such a conclusion is not new to psychological literature, it is significant that recent carefully worked-out studies substantiate earlier opinions in this regard. It should again be noted that the present conclusions are restricted to over-all measurements of mental status. Similar comparisons on more specific types of performance may, and in fact do, reveal systematic differences between the sexes.

TABLE II. — SEX DIFFERENCES ON THREE NON-VERBAL TESTS *

Investigator and Date	*Group*	*Test*		*Age, Years*										
				6	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
Pintner (11) 1924	Standardization Group (?)	Pintner Non-language Group Test	Boys					469		609				
			Girls					455		737				
			CR_M					−.5		.6				
			$CR\sigma$					−1.8		−.8				
Goodenough (3) 1926	A Sampling of California and Southern White Children	Goodenough Draw-a-Man Test	Boys	136	155	154	161	138	87					
			Girls	133	138	146	140	111	87					
			CR_M	−1.9	−1.6	−2.4	−1.5	−2.7	−.5					
			$CR\sigma$	.9	−1.1	1.7	.0	1.6	.8					
Lincoln (7) 1927	Children of Grade II to High School in 3 Eastern Communities	Dearborn General Examination	Boys		110	190	212	194	205	202	219	171	121	98
			Girls		123	204	210	192	190	190	175	160	134	122
			CR_M		1.9	−2.0	.4	1.7	.6	1.3	1.2	−.7	1.5	.5
			$CR\sigma$		−2.0	−.6	2.0	2.0	.0	.5	.0	−1.3	−1.0	1.0

* See Table I for further information.

REFERENCES

(1) COMMINS, W. D. "More about sex differences" *School and Soc.*, 28: 1928, 599–600.

(2) CONRAD, H. S., JONES, H. E., and HSIAO, H. H. "Sex differences in mental growth and decline." *Jour. Educ. Psychol.*, 24: 1933, 161–169.

(3) GOODENOUGH, F. L. *Measurement of Intelligence by Drawings.* (World Book Co.: Yonkers, N. Y., 1926) 173 pp.

(4) GOODENOUGH, F. L. "The consistency of sex differences in mental traits at various ages." *Psychol. Rev.*, 34: 1927, 440–462.

(5) HALES, N. M. *An Advanced Test of General Intelligence: A Report on the Revision, Extension and Standardization of a Form of the American Army Examination Alpha.* (Melbourne University Press: Melbourne, Australia, 1932) 60 pp. (Melbourne Univ. Education Research Series, No. 11.)

(6) HEILMAN, J. D. "Sex differences in intellectual abilities." *Jour. Educ. Psychol.*, 24: 1933, 47–62.

(7) LINCOLN, E. A. *Sex Differences in the Growth of American School Children.* (Warwick and York: Baltimore, 1927) 189 pp.

(8) MCNEMAR, Q., and TERMAN, L. M. "Sex differences in variational tendency." *Genet. Psychol. Monog.*, 18: 1936, 1–66.

(9) MARSHALL, A. J. "The standardization of Spearman's 'measure of intelligence' for Perth, Western Australia." *Austral. Council Educ. Res. Ser.*, No. 22: 1934, 23–50.

(10) MILES, C. C. "Sex in social psychology." [In] *Handbook of Social Psychology.* Ed. C. Murchison. (Clark University Press: Worcester, Mass., 1935) 1195 pp.

(11) PINTNER, R. "Results obtained with the non-language group test." *Jour. Educ. Psychol.*, 15: 1924, 473–483.

(12) PRESSEY, L. W. "Sex differences shown by 2,544 children on a group scale of intelligence, with special reference to variability." *Jour. Applied Psychol.*, 2: 1918, 323–340.

(13) ROBERTS, J. A. F., NORMAN, R. M., and GRIFFITHS, R. "Studies on a child population. I. Definition of the sample." *Annals of Eugen.*, 6: 1935, 319–338.

(14) SCHILLER, B. "Verbal, numerical and spatial abilities of young children." *Arch. Psychol.*, 24: 1933–1934, pp. 69.

(15) SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION. *The Intelligence of Scottish Children: A National Survey of an Age-Group.* (University of London Press: London, 1933) 160 pp.

(16) TERMAN, L. M. *The Stanford Revision and Extension of the Binet-Simon Scale for Measuring Intelligence.* (Warwick & York: Baltimore, Md., 1917) 179 pp.

(17) THORNDIKE, E. L. "Sex differences in status and gain in intelligence scores from thirteen to eighteen." *Ped. Sem. and Jour. Genet. Psychol.*, 33: 1926, 167–181.

(18) Wellman, B. L. "Sex differences." [In] *Handbook of Child Psychology.* Ed. C. Murchison, pp. 626–649. (Clark University Press: Worcester, Mass.) 956 pp.

(19) Whipple, G. M. "Sex differences in intelligence-test scores in the elementary school." *Jour. Educ. Res.*, 15: 1927, 111–117.

(20) Whitmire, E. D. *A Study of Sex Differences in 1340 Unselected Children.* (Stanford University, Unpublished Master's Thesis, 1920)

CHAPTER VII

EFFECTS OF HUMAN FERTILITY TRENDS UPON THE DISTRIBUTION OF INTELLIGENCE AND CULTURE

Raymond B. Cattell
Professor of Genetic Psychology
Clark University
Worcester, Massachusetts

I. The Problem of Biological Changes in the Constitution of Population

Without reasonably exact calculations as to the resources of intelligence available in the population of tomorrow the plans of educators are likely to be so much scrap paper. For, firstly, schools and curricula must be adapted to pupils' abilities; and, secondly, the heritage of native mental capacity will shape the social ideas of the community for which education is a preparation (10).

Available intelligence in any community may be affected by selective migration, by extent of opportunities for social promotion, and by differential death and birth rates. There is reason to believe that some of the major historical changes, including the rise and fall of cultures, have been rooted in these factors (36).

Under present civilized conditions the differential birth rate towers above the other factors. As Fisher concludes (21): " The intensity of selection by differences of fertility is [in man] relatively enormous in comparison to selective intensities to be expected in nature."

But it is only in the last ten years that techniques have become available that will enable us to find out how differential reproduction is likely to affect reserves of intelligence, or to analyze the interaction of causes and effects in this process. The techniques depend on more precise conceptions in population statistics, more knowledge regarding the inheritance of intelligence, and more controlled large-scale surveys of intelligence-fertility data.

Nevertheless, most observers have definitely taken the view, since the beginning of this century, that the average intelligence has been declining during that period. They have had to deduce the drift from such straws of evidence as (*a*) that recognized mental deficiency has increased (26), (*b*) that schools and clinics report impressively large families among the dull and backward, whereas 'only children' are more prevalent among the highly gifted, and (*c*) that, considered from the standpoint of social status, the higher strata consistently show lower birth rates than the lower strata.

This ill-assorted evidence we will consider among the 'indirect' researches, reserving our main reasoning for the *ad hoc* investigations. The indirect evidence from the pre-intelligence-test era (5, 20, 23, 28) has been reviewed by Dawson (17) and points almost uniformly to an inverse relation between amount of intelligence and amount of reproduction. The further data now available may be briefly surveyed as follows.

II. The Evidence on Differential Fertility and Intelligence

Indirect evidence, which depends on comparisons of fertility and intelligence between certain natural, largely inbreeding groups, considered as a whole, includes such genetically differentiated groups as social classes, races, geographical localities, rural and urban divisions, occupational classes, and other functional groups.

Burt and Terman noted twenty years ago that better intelligence scores were made by children from the upper and middle classes. Owing to the heat of political feelings, these findings have been much assailed, but rarely by scientists (41), and numerous researches, with varied tests and with children reared apart from their parents (15, 18, 27, 30, 31, 32, 35, 43, 48, 50), leave no doubt as to the reality of some biological class differentiation. A shrewd review of researches available to 1933 is made by Lorimer and Osborn (34) who settle on an average intelligence-status correlation of +0.3.

Through occupational and social surveys—for example, in the service of vocational guidance—it is known that there is a parallel, but closer correlation of social status and parental (adult) intelligence (6, 22, 34). A closer association is to be expected, since, owing to 'regression to the mean' and other factors, the intelligence of children is more central in tendency (with respect to the community) and, in larger families, more variable about the family mean than that of their parents. The relation of social status to fertility has been established, through the work of sociologists, for a much longer period and no less precisely (14, 25, 28, 34, 36, 53). It has been an inverse relation since at least 1870 in Britain and 1890 in America.

Other middle terms in the syllogism of indirect argument are less universal and therefore less important. Some races are characterized in certain countries by higher fertility and lower average intelligence than others. Southern Europeans, American Indians, Chinese, and Mexicans have unusually high fertility in the United States (34). Maller (38) gives the following from his adequately extensive survey of two such groups.[1]

Areas	*Mean IQ*	*Births per 1000*
15 Jewish	106.2	18.45
15 Italian	87.8	23.15

When an allowance for environmental handicap in test performances has been made, rural groups, both in America and Britain (7, 34) average about 5 points of IQ lower than city groups. Their fertility is greater.

The school districts of New York City, when analyzed for intelligence of school children and average size of family in the schools, show a correlation of —0.34 between these variables (37), and there is an indication of a similar trend when states of the union are taken as units (34).

One difficulty here is that none of these forms of indirect evidence permit of a superstructure of calculation or even of the logical certainty of a relation between intelligence and fertility, for a correlation of *A* with *B* and of *B* with *C* does not prove a correlation of *A* with *C*. Nevertheless, as the correlations become higher, and since the trend is consistent, only very special circumstances would negate the probability of such an association. A second difficulty is that the mean difference of intelligence-test scores associated with social classes cannot be taken at its face value. Part of it is undoubtedly due to environmental susceptibility in the tests.

Direct evidence was first brought forward in 1923 by Duff and Thomson (19) in England, and in 1924 by Haggerty and Nash in America (27). The American results were extended in 1925 by Chapman and Wiggins (13) who obtained a correlation of —0.33 (650 cases) between child's intelligence and size of family (fraternity); by Terman, 1926 (54), with an r of -0.27 ± 0.06 (92 cases); by Conrad and Jones (16); by Willoughby (56); and finally by the large-scale research of Lentz (33) in 1927 yielding $r = -0.304$ (4,330 cases). With

[1] See also Reference 1.

the exception of two researches (16, 56), which tested parents[1] and consequently dealt with small and selected numbers (less than 100 pairs), the results consistently center upon a correlation of approximately —0.3.

The regional sampling in Britain has been most satisfactorily widespread. Bradford (2) confirmed the first results in 1925 with an r of —0.25 (393 cases). Dawson (17) in 1932 added Scottish results, obtaining an r of -0.19 ± 0.02 with individual testing on a population of 1,239.

In spite of variations in regional sampling and social range (Dawson's group was largely laborers; Terman's largely business and professional), both the direction and magnitude of the correlations are surprisingly uniform. Indeed, with the aid of later results, it is possible even to detect a consistent slight difference in the trend of the correlations as between English and American populations, the former averaging about —0.2, the latter about —0.3.

The trend in the British population was put beyond doubt by two recent large-scale researches taking special precautions to get complete sampling of the communities covered. The present writer, using a special non-verbal ' G ' test (7) to avoid cultural effects, took a cross section of one typical industrial city and one rural region (3,734 families). Fraser Roberts (46, 47) worked similarly with an old provincial city (3,361 families), obtaining an r of —0.224. The British results to 1937 are appraised and summarized by Bradford (3). It is desirable that they now be supplemented by tests on parents as well as children.

III. Analysis of Natural Selection with Respect to Intelligence

Calculation of changes in the nation's distribution of intelligence resulting from these differential birth rates requires attention to a variety of biological and social factors. Although a digression into detailed arguments concerning the inheritance of intelligence is impossible here, its importance to the issue requires a statement of the general conclusions accepted as a foundation. Probably the most valuable single source of evidence is a previous Yearbook (40). The

[1] The present writer saw no reason to conclude that it is other than a coincidence, or a consequence of selection associated with small samples, that the surveys using parents rather than children yield the only non-significant correlations.

nature-nurture problem has been summarized to 1933 by Schwesinger (49), and less exhaustively, by the present writer up to 1936 (7). Fresh evidence is given in the present Yearbook.

Most psychological opinion concurs with the conclusion of the classically thorough research of Burks (4); namely, that 80 percent of IQ variance is due to heredity and 20 percent to environment. I believe that the interpretation assigns too much to environment, neglecting the presence of environmental skills in the test itself, for most of the tests used in nature-nurture inquiries, including the Binet, are demonstrably weighted unnecessarily with cultural elements (8, 11), with the result that investigators find 'environmental influence on intelligence' to the extent to which they first admit it to the test. This seems the most reasonable explanation of some recent anomalous results with very young children (11).

With somewhat more culture-free tests I found a mid-parent-mid-child correlation of 0.91 ± 0.028, and argued (12) that lower previous figures are due to selection (notably deficient variability), lack of correction for attenuation, of age correction in parents, and so forth. This would indicate that 80 percent of child interfamilial variance in IQ is determined by the intelligence of the parents. Because of segregation (that is, because the immediate parents do not alone determine the inheritance) the inheritance element would then be somewhat above 80 percent.

Consequently, in constructing the hypothetical frequency polygon of the next generation little error is introduced through multiplying the number of people in each intelligence interval by the net reproduction rate found in that interval. For, in any case, there is an 'inheritance' of family attitude almost as persistent as biological inheritance, which would contribute to determining the 'effective intelligence' of the next generation in the manner here assumed.

Figures I and II show, respectively, the family sizes found for each intelligence interval in the writer's English research and the results of applying the family size as a factor to each intelligence interval (whilst dividing throughout by a constant required to keep the population stationary; since in fact it is stationary — the discrepancy between reproduction rate and population growth arising from influences now to be considered).

The decline of the mean by 3.1 points is considered only a rough approximation, since, in exact calculation, the following nine influences are also to be taken into account:

1. A differential celibacy rate.

2. A differential ' barren marriage ' rate. The magnitude of these first two influences, in their non-differential aspect, can be gauged from the findings of population statisticians (8, 14) that of 1,000 girl children born, between 600 and 700 marry and that about 1 in 6 of the marriages is completely sterile.

3. A differential death rate. This is operative as it applies to the prereproductive period only. It is in the opposite direction to the birth rate, but by

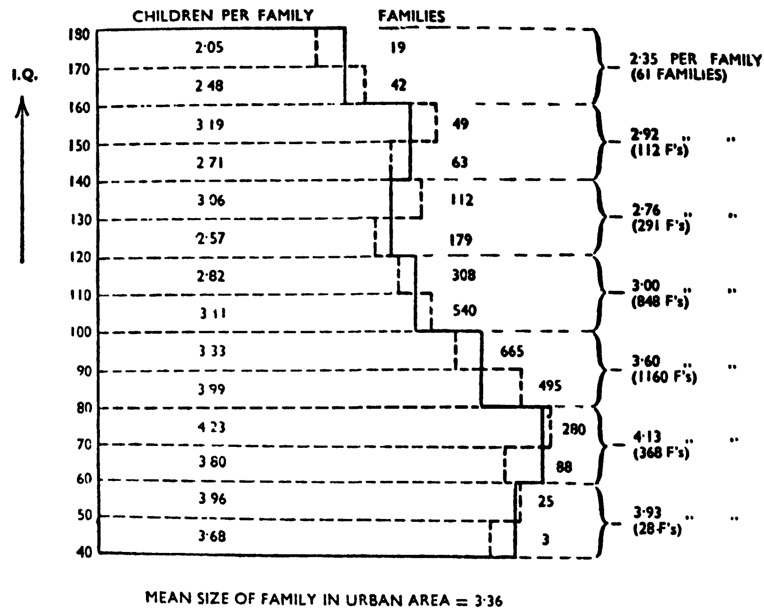

FIG. I. — URBAN AREA: SIZE OF FAMILY AND INTELLIGENCE QUOTIENT.

no means nullifies it. Indeed, we may expect that with progress in welfare work it will practically cease to modify the results of the differential birth rate.

4. Differences in the customary age of reproduction (length of generation).

5. The presence of incomplete families in the material gathered. Most studies have dealt with ten- to twelve-year-old children. Those families in which the representatives happened to be the oldest child might not be complete, but the product of two such probabilities must obviously be small. Opinion has favored the view that, if families were complete, the negative correlation of fertility and IQ would be augmented (small families being more likely to be complete), and Dawson experimentally showed this to be true in his samples. Conrad and Jones (16), however, give an interesting illustration of the opposite effect in their small sample of rural New Englanders.

6. The extent of assortative mating in the community as a whole. In so far as mating is unassortative, the fertility factor properly to be applied may be that of the more or less intelligent parent rather than their mean. Fisher's work (21) suggests that the mother's cultural status is more important in determining fertility.

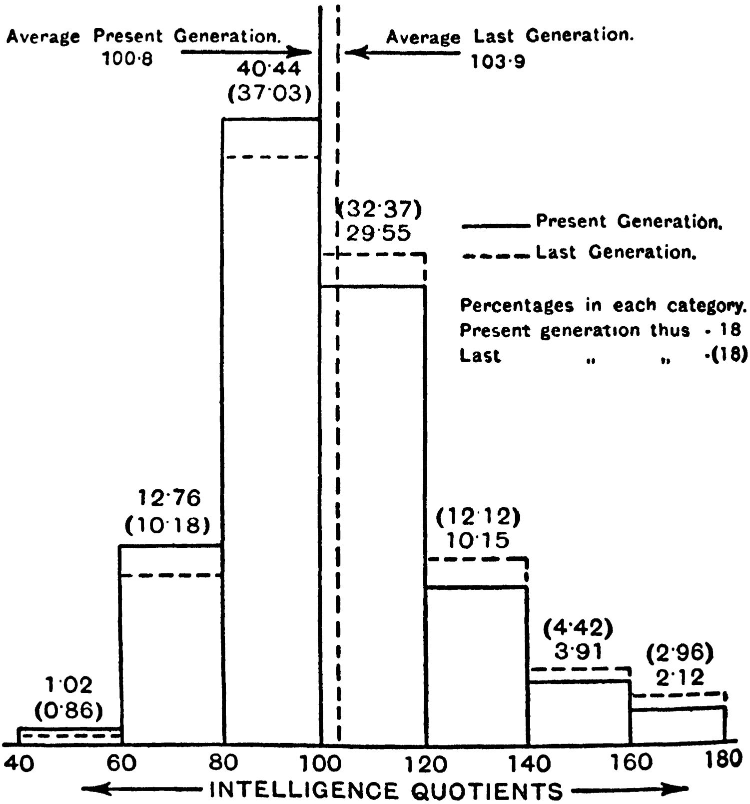

FIG. II. — ALTERATION OF DISTRIBUTION OF INTELLIGENCE. PRESENT ADULTS TO PRESENT TEN-YEAR-OLDS (1936). URBAN AREA.

7. Hereditary 'Regression to the Mean' (Galton's law). Various researches show children to have only one-half to two-thirds of the parental deviation from the intelligence mean, and vice versa, depending upon the extent of Factor 6 and of social segregation in marriage — inbreeding tending more to a 'true line.' One way of considering the necessary correction for regression is to suppose that the reproduction rate should properly be applied to an interval nearer the mean than that for which it is discovered. In addition to regression, there would be greater dispersion of IQ in the offspring

generation. Both of the corrections for these effects would decrease the apparent intergeneration IQ change.

8. The degree of epistasis. As Willoughby (57) has shown in an incisive analysis, a highly exact calculation would need to take into account the amount of epistasis in inheritance due to the action of dominants in the presumably multi-gene determination of intelligence.

9. The extension of the fertility factor. Since the intelligence of siblings will be both regressive and dispersed relative to that of the sample child taken as a representative, the application of the fertility factor should again be extended outside the category to which it is applied (as in 7); this again would lessen the decline as calculated by the rough method.

No research has applied these corrections, nor are the necessary data available on many points. Since they are small and about equally balanced in tendency — Factors 1, 2, 4 and possibly 5 exaggerating the trend and Factors 3, 6, 9 and possibly 8 reducing it — the approximate calculation made in this first approach is, we feel, reasonably dependable. The factors have been listed to enable the reader to form his own opinion on such debatable new ground.

IV. Calculation of a One-Generation Change

Three calculations of change in community IQ are now available: those by Lentz (33), by Maller (38), and by the writer (7). The first study and the last one use the same form of data and calculation and yield a decline of mean IQ per generation of 4.4 in the American results and 3.0 in the British results. It is interesting to note that the difference accords with the difference of correlation coefficients. Maller's calculation, as he himself points out, is inadequate, being based on factors applied to school districts rather than families, and being on a population (New York City) not representative of America. His figure is almost exactly the same as that for Britain — 3.0 points.

The inverse relation of IQ and fertility and its intensity are local in space and time. It is generally agreed that a differential survival rate in the present direction did not exist in America much before the end, or in Britain before the middle, of the last century. The rate varies with the cultural and economic levels of a country and with different forms of social organization. There is evidence of a similar differential birth rate in Scandinavia and Germany, and inferentially in the British Dominions, France, Italy, Russia, and India. Elsewhere there seem few statistics even for an inferential statement as to population trends in IQ.

So radical a change as the ironing out of social-status differences does not alone suffice to remove the trend, for the observations of Gantt (24) and the test survey of Price (44) show that it still exists in Russia. This is in accordance with expectations from recent analyses here, which show that fertility is more closely associated with intelligence than with social or cultural status (34, 38); and that the negative correlation with intelligence exists even within any one occupation (9). Small groups — university alumni, the professions, the nobility — however, occasionally show a eugenic trend within themselves (9, 34). But, except for the nobility, their birth rate as a whole is not sufficient to maintain their numbers.

Recently in social groups with a developed sense of social responsibility and lacking racial strata of subcultural intelligence level, a positive correlation of intelligence with fertility has occasionally appeared. An incomplete and unpublished research by Willoughby into the families of New England high-school graduates agrees with the Conrad and Jones New England material in showing at least an absence of negative relationship.

These recent results perhaps begin to show the results of a generation's spread of birth control practice into the lower intelligence groups. In Scandinavia and notably in Stockholm (7) there is reason to believe that the dysgenic trend has now been halted for almost a generation.

V. Causes and Consequences of the Differential Birth Rate

It will be evident that the causes of the differential birth rate are both simple and complex. The absence of planned contraception among the less intelligent is undoubtedly the main cause; but the reasons for excessive birth limitation among the more educated are subtle and not fully understood, being more psychological than crudely economic, in spite of popular assumptions to the contrary (7). As Fisher (21) and Wagner-Manslau (55) have shown, the birth rate is adequate in the nobilities, those of assured social status, and those with no aspirations to social status (the indigent poor or the so-called 'depressed' classes); that is, those classes that can neither rise nor fall socially. Excessive limitation of families is found in the upper and the lower middle classes, and especially in caste-free populations (21) readily permitting social promotion. Perhaps that is one causal factor in the differential birth rate being steeper in America than in Britain.

Any attempt to work out the consequences of a fall of the IQ by

1.0 to 1.5 points per decade (as found in these studies), if that fall should persist for several decades, involves some speculation, both on account of inadequate principles in social psychology and of our inability to predict changes in other relevant factors. Elsewhere (10) reasons have been set out indicating the likelihood of the following changes. They presuppose that a lowering of the mean will be accompanied by an increase in the standard deviation, due to the uppermost groups maintaining their numbers more in the near future.

1. A fall of scholastic standards in 'G'-saturated school subjects (mathematics, English),
2. An increase in delinquency,
3. An increase of unemployment at the unskilled level, and indirectly therefrom, less markedly, at all levels,
4. An increased conservatism, shown, for example, in a reversion to more rigid religious forms; or, at least, in an arrest of the development out of such forms,
5. An increased susceptibility to propaganda,
6. A shift from democratic toward autocratic or bureaucratic government.

REFERENCES

(1) ARSENIAN, S. "The Spearman Visual Perception Test (Part I) with pantomime directions." *Brit. Jour. Educ. Psychol.*, 7: 1937, 287–301.

(2) BRADFORD, E. J. G. "Can present scholastic standards be maintained?" *Forum Educ.*, 3: 1925, 186–198.

(3) BRADFORD, E. J. G. "The relation of intelligence to varying birth rate in different social grades." *Brit. Jour. Educ. Psychol.*, 7: 1937, 229–246.

(4) BURKS, B. S. "Nature and Nurture: Their Influence upon Intelligence." *Twenty-Seventh Yearbook, Part I,* of this Society, 1928, pp. 219–316.

(5) CATTELL, J. McK. "Families of American men of science." *Scientific Monthly,* 4: 1917, 248–262; 5: 368–377.

(6) CATTELL, R. B. "Occupational norms of intelligence and the standardization of an adult intelligence test." *Brit. Jour. Psychol.*, 25: 1934–1935, 1–28.

(7) CATTELL, R. B. *The Fight for Our National Intelligence.* (P. S. King: London, 1937) 164 pp.

(8) CATTELL, R. B. "Measurement versus intuition in applied psychology." *Character and Personality,* 6: 1937, 114–131.

(9) CATTELL, R. B. "Some further relations between intelligence, fertility, and socio-economic factors." *Eugen. Rev.*, 29: 1937, 171–179.

(10) CATTELL, R. B. "Some changes in social life in a community with a falling intelligence quotient." *Brit. Jour. Psychol.*, 28: 1938, 430–450.

(11) CATTELL, R. B. "A culture-free intelligence test." *Jour. Educ. Psychol.*, 1939.

(12) CATTELL, R. B., and WILLSON, J. H. "Contributions concerning mental inheritance. I. Of Intelligence." *Brit. Jour. Educ. Psychol.*, 8: 1938, 129–149.

(13) CHAPMAN, J. C., and WIGGINS, D. M. "Relation of size of family to intelli-

gence of offspring and socio-economic status of family." *Ped. Sem.*, 32: 1925, 414–421.

(14) Charles, E. *The Twilight of Parenthood.* (Watts: London, 1934) 266 pp.

(15) Collins, J. E. "The intelligence of school children and parental occupation." *Jour. Educ. Res.*, 17: 1928, 157–169.

(16) Conrad, H. S., and Jones, H. E. "A field study of the differential birth rate." *Jour. Amer. Statis. Assoc.*, 27: 1932, 153–159.

(17) Dawson, S. "Intelligence and fertility." *Brit. Jour. Psychol.*, 23: 1932–1933, 42–51.

(18) Dexter, E. S. "The relation between occupation of parent and intelligence of school children." *School and Soc.*, 17: 1923, 612–614.

(19) Duff, J. F., and Thomson, G. H. "Social and geographical distribution of intelligence in Northumberland." *Brit. Jour. Psychol.*, 14: 1923–1924, 192–198.

(20) Ellis, H. H. *A Study of British Genius.* (Hurst: London, 1904) 300 pp.

(21) Fisher, R. A. *The Genetical Theory of Natural Selection.* (Clarendon Press: Oxford, 1930) Chaps. IX–XI.

(22) Fryer, D. "Occupational-intelligence standards." *School and Soc.*, 16: 1922, 273–277.

(23) Galton, F. *English Men of Science; Their Nature and Nurture.* (Macmillan: London, 1874) 270 pp.

(24) Gantt, W. H. *A Medical Review of Soviet Russia.* (Brit. Med. Assoc.: London, 1928) 112 pp.

(25) Great Britain, Census Office. *Census of England and Wales, 1911,* "Fertility of marriage." (H. M. Stationery Office: London, 1915) 424 pp.

(26) Great Britain, Mental Deficiency Committee. *Report of the Mental Deficiency Committee. Part IV.* "Investigation into the incidence of mental deficiency." Report by Lewis, E. O. (H. M. Stationery Office: London, 1929) 239 pp.

(27) Haggerty, M. E., and Nash, H. B. "Mental capacity of children and paternal occupation." *Jour. Educ. Psychol.*, 15: 1924, 559–572.

(28) Heron, E. *On the Relation of Fertility in Man to Social Status, and on the Changes in this Relation that Have Taken Place during Fifty Years.* (Dulau: London, 1906) 22 pp.

(29) Jones, D. C. "Eugenic aspects of the Merseyside survey." *Eugen. Rev.*, 28: 1936, 103–113.

(30) Jordan, A. M. "Parental occupation and children's intelligence scores." *Jour. Applied Psychol.*, 17: 1933, 103–119.

(31) Kornhauser, A. W. "Economic standing of parents and the intelligence of their children." *Jour. Educ. Psychol.*, 9: 1918, 159–164.

(32) Lawrence, E. M. "An investigation into the relation between intelligence and inheritance." *Brit. Jour. Psychol., Monog. Suppl.*, No. 16, 5: 1931. 80 pp.

(33) Lentz, T. F., Jr. "Relation of I.Q. to size of family." *Jour. Educ. Psychol.*, 18: 1927, 486–496.

(34) Lorimer, F., and Osborn, F. *Dynamics of Population: Social and Biological Significance of Changing Birth Rates in the United States.* (Macmillan: New York, 1934) 461 pp.

(35) MACDONALD, H. "Social distribution of intelligence in the Isle of Wight." *Brit. Jour. Psychol.*, 16: 1925–1926, 123–129.

(36) MACDOUGALL, W. *National Welfare and National Decay.* (Methuen & Co.: London, 1921) 213 pp.

(37) MALLER, J. B. "Mental ability and its relation to physical health and socio-economic status." *Psychol. Clin.*, 22: 1933, 101–107.

(38) MALLER, J. B. "Vital indices and their relation to psychological and social factors." *Human Biol.*, 5: 1933, 94–121.

(39) MOORE, E. "Report of the Work of Commission II: (Differential fertility)." [In] *Problems of Population. Report of the Proceedings of the Second General Assembly of the International Union for the Scientific Investigation of Population Problems.* (George Allen and Unwin, Ltd.: London, 1932) pp. 355–357.

(40) NATIONAL SOCIETY FOR THE STUDY OF EDUCATION. *The Twenty-Seventh Yearbook. Part II.* "Nature and Nurture: Their Influence upon Intelligence." (Public School Publishing Co.: Bloomington, Ill., 1928) 397 pp.

(41) NEFF, W. S. "Socio-economic status and intelligence: A critical survey." *Psychol. Bull.*, 35: 1938, 727–757.

(42) PEARL, R. *The Biology of Population Growth.* (Knopf: New York, 1925) 260 pp.

(43) PRESSEY, S. L., and RALSTON, R. "Minor studies from the psychological laboratory of Indiana University: The relation of general intelligence of school children to the occupation of their fathers." *Jour. Applied Psychol.*, 3: 1919, 366–373.

(44) PRICE, B. "Reference data in Moscow families of 1935." (Abstract) *Psychol. Bull.*, 35: 1938, 696–697.

(45) ROBERTS, J. A. F. "Intelligence and family size." *Eugen. Rev.*, 30: 1939, 237–247.

(46) ROBERTS, J. A. F., NORMAN, R. M., and GRIFFITHS, R. "Studies on a child population. I. Definition of sample." *Annals of Eugen.*, 6: 1935, 319–388.

(47) ROBERTS, J. A. F., NORMAN, R. M., and GRIFFITHS, R. "Studies on a child population. III. Intelligence and family size." *Annals of Eugen.*, 8: 1938, 178–215.

(48) SANDIFORD, S. "Parental occupation and the intelligence of offspring." *School and Soc.*, 23: 1926, 117–119.

(49) SCHWESINGER, G. C. *Heredity and Environment. Studies in the Genesis of Psychological Characteristics.* (Macmillan: New York, 1933) 484 pp.

(50) STROUD, J. B. "A study of the relation of intelligence-test scores of public-school children to the economic status of their parents." *Ped. Sem. and Jour. Genet. Psychol.*, 35: 1928, 105–111.

(51) SUTHERLAND, H. E. G. "The relationship between I.Q. and size of family." *Jour. Educ. Psychol.*, 20: 1929, 81–90.

(52) SUTHERLAND, H. E. G., and THOMSON, G. H. "The correlation between intelligence and size of family." *Brit. Jour. Psychol.* (Gen. Sect.), 17: 1926–1927, 81–92.

(53) SYDENSTRICKER, E., and NOTESTEIN, F. W. "Differential fertility according to social class." *Jour. Amer. Statis. Assoc.*, 25: 1930, 9–32.

(54) TERMAN, L. M., BALDWIN, B. T., BRONSON, E. T., and OTHERS. *Genetic*

Studies of Genius. I. Mental and Physical Traits of a Thousand Gifted Children. (Stanford University Press: Stanford, 1925) 648 pp.

(55) WAGNER-MANSLAU, W. *Ber. d. Internat. Kong. f. Bevolkerungswissen.* 1933.

(56) WILLOUGHBY, R. R. "Fertility and parental intelligence." *Amer. Jour. Psychol.*, 40: 1928, 671–672.

(57) WILLOUGHBY, R. R., and GOODRIE, M. "Neglected factors in the differential birth rate problem." *Ped. Sem. and Jour. Genet. Psychol.*, 34: 1927, 373–393.

CHAPTER VIII

TEN YEARS OF RESEARCH ON TWINS: CONTRIBUTIONS TO THE NATURE–NURTURE PROBLEM [1]

Harold D. Carter
Research Associate, Institute of Child Welfare
University of California
Berkeley, California

I. Introduction

Many persons have been disappointed by the results of twin studies in which solutions to nature-nurture problems have been sought. There are several reasons. One is the prevalence of studies unequal to the requirements of the problems attacked. Another reason is the tendency of critics to expect too much, either because of lack of appreciation of the complexity of nature-nurture problems or because of failure to realize the limitations of the twin-study technique. This review of methods and results aims not only to indicate some trends of recent contributions, but also to present a picture of the inherent difficulties and complexities of the problems.

This review will not be exhaustive in any sense. In so far as is practicable, it will deal critically with methods and results rather than with particular investigations. The material may be summarized under the following topics: classification of twins, abnormalities in twins, studies of twins reared apart, intensive studies of twins reared together, twin resemblances in relation to age, correlational studies of twin resemblances, and environmental influences. Each of these will be discussed in a separate section; in each such discussion a minimum of space will be devoted to the elementary facts about twins known prior to 1928. The main emphasis is upon trends in twin studies, and advances in knowledge since the publication of the *Twenty-Seventh Yearbook* of this Society.

II. Classification of Twins

The classification of twins into monozygotic (identical) and dizygotic (fraternal) groups is basic to the whole approach; the last word

[1] Clerical assistance in connection with this report was furnished by the personnel of Works Progress Administration Official Project No. 465–03–3–61.

has not yet been written on problems of diagnosis, but important contributions have been made in recent research. Studies using fairly large samples (1, 15, 76, 133) usually show that in only two or three cases in a hundred is there difficulty in arriving at a dependable diagnosis. The difficulties, we find, arise largely through the operation of three factors: first, inexperience of the worker, which is particularly significant where several racial groups are represented; second, the great similarity of siblings in some families; third, the fact that identical twins *may* differ very greatly in size and appearance as the result of accident, or developmental abnormality, or unusual mirror-imaged asymmetry. The difficulties are not very serious so far as psychological studies are concerned, because the incidence of troublesome cases is too small to invalidate statistical studies of large numbers, and because the logic of intensive studies of small numbers does not usually require inclusion of the doubtful cases in the experiment.

Palm-print and fingerprint material has been shown (79, 80, 96) to be a valuable aid in diagnosis. It is definitely established, however, that the dermatoglyphic method is unreliable in individual cases (7, 19, 92, 96). A number of studies (20, 51, 118, 126) have shown that the obstetric method of diagnosis on the basis of fetal membranes is also a fallible technique. These methods are now being regarded as objective aids mainly useful in conjunction with other methods. Rife (94, 96) has shown that the pattern of iris pigmentation is a particularly valuable item in diagnosis, as are, for older subjects, the presence and distribution of hair on the hands. The use of blood-agglutination groups has proved valuable (94, 96, 110, 117). Schiller (111) has also shown that microscopic capillary diagnosis can be helpful. Recent research indicates that the best procedure in diagnosis is that which makes use of a variety of items; no one of the inherited anatomical characteristics is adequate taken alone, but a sufficient number of them establish the probability of correct diagnosis beyond reasonable doubt in most instances. The general impression of similarity, based upon inspection, is a valuable subjective extension of this resemblance-probability method.

III. Abnormalities in Twins

Most of the evidence concerning abnormalities in twins consists of case studies. These cover a wide range of topics. They usually show that identical twins are more similar than fraternal twins or siblings in physical defects, in criminality, in disease histories, in psychoses, and in the specific symptoms, onset, and development of diseases. In visual handicaps, specifically in refractive errors, twins are more alike than siblings (54, 132). Identical twins, more often than fraternal twins, are alike in the tendency to have tuberculosis (22, 56, 90), diabetes (51, 128), and syphilis (88).

Sometimes one twin suffers from birth injury, resulting in physical defect and accompanying mental defect. Thus Lewis (58) has published a report of identical twins who are extremely dissimilar in physique because of acromegaly in one member of the pair. But in cases of schizophrenia or of epilepsy in one-egg twins, we may expect both twins to be affected (6, 53, 97, 103). According to case reports, the onset of diseases is often simultaneous in identical twins, and the specific symptoms and general course of the disease reveal striking similarities. With respect to most diseases reported, there are exceptions, especially in ailments apparently resulting from glandular disturbance, such as hypothyroidism or acromegaly.

Numerous reports attest to a similarity in the psychoses of identical twins. In cases of schizophrenia (33) and of manic-depressive insanity (4) the similarity in identical twins often extends to the specific symptoms manifested. According to the statistical surveys available, in all varieties of mental disorders studied, both members of a pair of monozygotic twins are usually affected, whereas in fraternal twins it is more often only one who is affected. This difference between the two sorts of twins has been reported for behavior problems, crime, delinquency, mental deficiency, epilepsy, dementia praecox, and manic-depressive insanity (101, 102). Comparison of the twin data with similar data on siblings (41) seems to indicate that the similarity of disorders parallels the genetic relationship.

In general health and in the incidence and severity of children's diseases, monozygotic twins are more similar than dizygotic twins (56, 132). These conditions of general health are, of course, factors that probably affect intellectual stimulation and activity as well as physical growth and well-being.

Studies of feeblemindedness in twins show that in the cases of identical twins usually both are feebleminded (101, 115, 116). The work of Rosanoff, Handy, and Plesset (101) indicates that heredity is an important, but not a necessary or adequate, explanation of feeblemindedness; these authors conclude that cerebral birth trauma is very important in the etiology of mental deficiency.

Some case studies (60) indicate the probable heritability of specific mental defects, such as alexia and agraphia. It is, of course, obvious that any hereditary abnormality that interferes with tool skills, like reading, is an important factor in mental development and in achievement.

Survey of the literature shows that in all reported cases of mongolism in identical twins both twins have always been affected, whereas of all reported cases of mongolism in fraternal twins, only one has ever been affected (30, 34, 52, 99). An equally marked difference in incidence has not been reported in other characters; possibly exceptions would have been found here if a diagnosis of zygosity had always been available when mongolism was reported. The evidence tends to discount explanations of mongolism based upon vague theories as to the inadequacy of intra-uterine environment, or the general condition of the mother. The data from studies of twins suggest either that mongolism is inherited or that it involves anomalous development at a very early germinal stage.

These recent studies of the incidence of mental and physical disorders in twins seem to aid in determination of the causal factors and in distinguishing between hereditary and environmental factors in mental development. It is still not entirely clear whether certain disorders have an inherited basis or are induced by environmental stimulation, but studies of twins are providing some factual basis for present-day theories. The study by Rosanoff, Handy, and Plesset (101), for example, in spite of its limitations, is one of the major contributions to our knowledge of the etiology of mental deficiency.

IV. Studies of Twins Reared Apart

The potential value of detailed studies of persons with identical heredity and different environment has long been recognized. To Newman, Freeman, and Holzinger (84) we are indebted for the largest step toward the realization of such potentialities. Their study includes 19 pairs of identical twins reared in separate homes. The differences in physique, intelligence, achievement, and personality between these twins are evaluated in the light of their environmental differences. Further interpretations are afforded by comparison with data from 50 pairs of identical twins and 52 pairs of fraternal twins reared together. A major conclusion is that physical traits are least affected by environment, that ability and achievement are more affected, and that traits of personality are most affected. Readers may well pay attention to the authors' further statements that the influences of nature and nurture constitute not one problem but rather a multitude of specific problems, and that there is no general solution of any major problem or even of any one of the specific problems at the present time.

Additional cases of twins reared apart have been reported by other authors. Bouterwek (3) has reported a case of twin sisters separated during the first nine years of life. These twins showed marked mirror-imaged asymmetry in physique. He suggests that mental differences in identical twins reared apart may sometimes be explained in terms of the asymmetry mechanism. Saudek (109) has reported a case of identical twins reared apart, in which intrapair differences were relatively small. He concluded that intelligence is modified by differences in environment much less than are emotional reactions.

The case studies reported by Newman, Freeman, and Holzinger (84) undoubtedly show that identical twins reared apart are more often different in physique, intelligence, and personality than are twins reared together. The evidence is not entirely consistent, however, for some of the twins reared apart were strikingly similar in spite of differences in environment. For example, in Case 7, one boy was reared on a farm and the other in town; there were no significant differences in intelligence. In Case 8, a difference of 15 points in IQ was found, although the difference in formal schooling was not great. The IQ of one twin was 92 and of the other was 77; apparently the one girl lived in a more stimulating social and cultural environment. The reader will wish to go over all the case histories, in order to reach his own conclusions as to the significance of the data.

The interpretation of differences between the IQ's of identical twins reared apart is not a particularly simple matter. Carter (13) has shown in intensive studies of three pairs of identical twins reared together that small but reliable differences in intelligence are revealed when the twins are repeatedly tested. Koch's study of a pair of Siamese twins showed that one was definitely superior to the other in IQ. Ley (59) has reported a case of identical twins reared together, in which the IQ's were 81 and 63. In one of Carter's cases, the differences in IQ were not in accord with the apparent environmental differences. Newman, Freeman, and Holzinger (84) are cautious in their conclusions. Their data show clearly that large environmental differences are accompanied by marked mental differences. In the majority of cases reported, the identical twins reared apart were not more different mentally than the above-mentioned cases of identical twins reared together. But in four or five cases in which the environments were greatly different, the IQ's of the twins were markedly different.

McNemar (69) has written a detailed criticism of the study by Newman, Freeman, and Holzinger, and a reply has been made by Holzinger (40), who admits the trenchancy of some of the criticisms, disagrees with others. Apparently, as is inevitable in any major study, some statistical errors were made. Taking account of appropriate statistical corrections for age and for range of talent increases the relative size of the coefficients of resemblance for identical twins reared apart, and makes them of an order of magnitude similar to those for identical twins reared together; the correlations for fraternal twins reared together are lower. It is clear from McNemar's paper and Holzinger's reply that the analysis and interpretation of data from studies of identical twins reared apart present many difficulties. Such data have offered no easy solution to nature-nurture problems.

The conclusion (84) that the effects of environment are greatest for personality traits, next for achievement, next for intelligence, and least for physique requires much qualification. Verschuer (123) and Newman (81) have shown that some physical traits are much more easily affected than others. McNemar (69) has pointed out that the measures of intelligence and personality traits used in the Newman-Freeman-Holzinger study have somewhat greater errors of measurement than the measures of physique that were employed. Allowance for the differences in trait variability in relation to age would produce a further reduction of the differences upon which the comparative statements were based. Some questions of experimental procedure further complicate this question. The conclusion seems reasonable, but it does not seem to have been clearly and quantitatively established. It is unreasonable to expect complete solution of such problems from data on 19 pairs of twins reared apart.

In summary, the studies of identical twins reared apart have shown clearly that drastic changes in the environment may be associated with corresponding mental differences to the extent of 15 or 20 points of IQ. Apparently, the cumulative effects of living for years in homes with different cultural stimulation may have a significant influence upon the IQ; also, it is evident that 4 or 5 years' difference in formal school training are usually associated with marked differences in achievement and in the manifestation of intelligence as measured by present-day tests. To the writer, the array of cases suggests also that the environmental differences between different homes (for example, town and country homes) are ordinarily less effective for the modification of

IQ's than many laymen and most proponents of nurture theories would like to believe.

V. Intensive Studies of Twins Reared Together

Under this heading are considered investigations in which twins were observed more intensively, or over a longer period of time, or with greater attention to detail than in the usual statistical survey. These investigations include observational case studies, experiments concerned with various aspects of reactivity, and training or learning experiments.

The studies of the Dionne quintuplets by Blatz (2) and his collaborators have been most detailed. The very comprehensive records of daily routine, training, and behavior should become increasingly valuable as time goes on. Such research should provide clinical insight into the factors underlying the various aspects of mental development. These reports and other available reports on triplets and quadruplets (8, 18, 51, 105) suggest that children of multiple births may be retarded mentally more often than single-born children. The quintuplets were retarded generally, and more in the development of speech functions than in other respects. These findings agree with the implications of recent studies of twins (10, 21, 29).

Studies of twins through the Rorschach method have shown somewhat greater similarity of identical than of fraternal twins (50, 125). The identical twins more frequently belong to the same Rorschach 'type' than do the fraternal twins; and the latter, more often than unrelated pairs. Yet the correlations between identical twins for various scores provided by this test are insignificant. This may arise from the incomplete validity and reduced reliability of the Rorschach method in normal samples.

A number of studies reveal a tendency of various responses of identical twins to be surprisingly similar in detail. Rife's study (95) indicates that they tend not only to get the same total scores on intelligence tests, but also to make the same errors on subtests. Carter's study (17) of inventory responses shows that identical twins tend not only to be more similar than fraternal twins in total scores for neurotic tendency, introversion, and the like, but the identical twins also tend to give the same answers to particular questions more often than do fraternal twins. Results of studies using the Rorschach test (50, 125) may be interpreted as indicating a similarity of identical twins in details of unidentified response mechanisms. Such material has been used as an indication of similarity of identical twins in pattern of response, or in qualitative aspects of response.

Several studies indicate the similarity of twins in various aspects of reactivity. Carmena (11), using a galvanometric method, found that twins are very similar in electrodermal responses. Carter (14) showed that the interests of identical twins are more similar than those of fraternal twins. Using a saliva-reflex method, Kanaev (47, 48) has shown great similarity in the development and course of conditioned responses in identical twins, in conditioned inhibitions, and even in the individual fluctuations of excitation. Hunt and Clarke (42) using a motion-picture method, found differences in the startle-reaction patterns of identical twins. Carmena (12), in a study of writing pressures using a kymograph-recording technique, showed striking similarity in the behavior of identical twins. The data of Carmena (12) suggest that character of movements and tonus of muscles play a part in determining the behavior of identical twins. All these studies are here regarded as interpretative and suggestive. They indicate similarities of identical twins in aspects of response that must undoubtedly affect the similarities in intellectual behavior, either directly, or through the operation of environmental influences.

The training or learning studies on twins have contributed little to the problem of nature-nurture in relation to intelligence; they have been largely concerned with the more overt aspects of behavior. Hilgard's study (36) by the method of co-twin control is an exception. This study is of interest here because of its inclusion of memory tests; that is, tests of higher mental functions. Practice given early in experimental period and practice given later were found to be equally effective as determinants of abilities retained after 3 or 6 months. This study indicates the greater importance of general developmental factors as compared with training. Strayer's study (121) of language development may be given a similar interpretation.

According to reports by Levit, Luria, Mirenova, and others (57, 63, 64, 65, 73, 74, 75), studies of twins in the U.S.S.R., using the co-twin control method, show that differences in training cause stable general differences in perception. Training in psychomotor functions is claimed to result in general improvement in such traits as activity, independence, and discipline. An effect of such training in increasing the IQ is also claimed. Luria (63) has argued that when mental functions are more complex, environment is more potent; whereas, when they are simple, heredity is more potent. Such an argument is hard to fit into the facts concerning effects of specific drill in school subjects, for it is very easy to modify simple mental functions, such as addition, and very difficult to modify complex functions, such as intelligence as now measured. In view of the fact that some of these findings are not in harmony with typical psychological investigations in this country, it is necessary to regard these new findings with caution until the experiments have been repeated in other laboratories.

VI. Twin Resemblances in Relation to Age

Various studies have shown clearly that twin resemblances in both mental and physical traits change with age. Weight may be taken as an example of an aspect of physique that reveals changes in twin resemblances with age in the course of development. Sontag and Nelson (118) have shown that early growth and nutritional condition were not the same at all times for a set of monozygotic triplets. Wilson and Jones (133) have shown that identical twins were not reliably more similar in birth weight than fraternal twins. As noted by Orel (85), weight differences change with age genotypically in fraternal twins, phenotypically in identical twins. Carter (13) has published a case study of identical twins who revealed no reliable weight differences in youth, but differed by 30 pounds at the age of 40.

The popular resemblance-probability method of diagnosis of monozygosity is known to be generally less dependable in infancy, because various aspects of physical individuality are less well developed in infancy. The diagnosis of the Dionne quintuplets (66) made much use of dermatoglyphics, in which twin resemblances do not change with age. It is also significant that diagnosis of monozygosity in twins is much easier during the school years than during maturity or middle age, because of changes in physique, muscular development, weight, coloring, appearance, and texture, color, and distribution of hair. The value of some studies has been lessened because of incorrect diagnosis in infancy. The frequency of doubtful diagnosis, or of failure to diagnose, is greater in infancy, and again in maturity.

We are not then necessarily committed to a belief in the greater modifiability of mental traits when we note that age changes take place in mental resemblances of twins. The tendency of several studies (36, 45, 121) is to show that during periods of rapid growth maturational factors seem much more important in the determination of twin resemblances than does specific training. As seen in the detailed studies by Blatz and others (2), with the increase in age different functions appear. Hence, age changes in resemblance must occur because of the fact that some functions are better measured at one age than at another. As McNemar (69) has pointed out, age changes in the range of trait variability and in the reliability of measurement techniques contribute to age changes in the magnitude of feasible measurements of twin resemblance. From the literature on physical growth one can find reason to expect lower correlation between measures of twins during periods of rapid growth, because of the instability of functions. Following the work of Richardson (93), one can also find a possible explanation of the fact that fraternal twins seem to be more alike in IQ than do siblings. Richardson reported that siblings tested when 3 years different in age were not more different than siblings tested when they were of the same age, but fraternal twins were more alike when tested at the same age than when tested two years apart. Carrying the argument further suggests that twins are more alike in the development of mental response mechanisms than are siblings, because the twins co-exist in time.

The earlier conception that environment should cause twins to get more alike as they grow older has long been considered illogical and has been dropped. There is reason to expect that, after the period of infancy, twins should become less alike as they grow older, and they are known to do so with respect to some few traits measured. Newman, Freeman, and Holzinger (84) have shown that fraternal twins become less alike in mental and physical traits as they grow older, whereas identical twins tend more nearly to maintain the same degree of resemblance. There are biological reasons for expecting this, as well as reasonable nurture hypotheses. Hence, findings concerning the age changes in resemblance are inconclusive in so far as the nature-nurture problem is concerned.

VII. Correlational Studies of Twin Resemblances

The recent tendency has been to seek material for nature-nurture arguments through comparison of twin resemblances in mental traits with those in physical traits that are supposedly less affected by environmental influences. This requires furnishing correlations obtained from other measurements, as a frame of reference for the similar data on IQ.

The correlations between identical twins in physical traits, such as height, cephalic index, and number of friction ridges, have been shown by Holzinger (38) to be in the neighborhood of .95, whereas similar correlations between fraternal twins are in the neighborhood of .60. Such evidence is usually employed as a standard of comparison for twin resemblances in IQ. The comparison is not completely satisfactory, since physical traits are used in diagnosis of zygosity and the mental traits are not so used.

Studies of the intelligence quotients of twins usually provide correlations between identical twins of about .85, and between fraternal twins of about .55 (9, 35, 38); figures approximating these values are found using either individual intelligence tests or group tests such as the Kuhlmann-Anderson, the Otis, or the Terman Group Tests. When specific abilities of twins have been studied, the correlations have usually been lower. For example, Brody (9) found correlations for mechanical ability to be .28 for fraternal twins and .69 for identical twins, using a test approximately as reliable as the above-mentioned group tests of mental ability. McNemar (68) found average coefficients of about .43 for fraternal twins and .79 for identical twins, using very reliable tests of motor skills. Higher correlations have been found for other measures of specific abilities, such as vocabulary, which are of the type frequently included in group tests of intelligence.

The literature on personality traits contains frequent implications that such traits are dependent upon environmental stimulation for their development. This is particularly true of vocational interests. In this field, Carter (14) found average correlations of .28 for fraternal twins and of .50 for identical twins, using 23 scales of the Strong Vocational Interest Blank. Using the Woodworth-Mathews test, Holzinger (38) obtained correlations for neurotic tendency of .37 for fraternal twins and .56 for identical twins. Using the Bernreuter Personality Inventory, Carter (15) reported the resemblance coefficients to be .32 for fraternal twins and .63 for identical twins, for neurotic tendency. For the six scales of the inventory, the resemblance coefficients varied about mean values of .27 for like-sex fraternal twins and .57 for identical twins. The scales used are at least as reliable as the group tests of intelligence used in the comparisons.

It appears that the correlations between twins in intelligence are of the same order of magnitude as those for physical measures if a slight allowance is made for the spurious resemblance in physique resulting from techniques of diagnosis. Slightly lower correlations are found when specific psychomotor functions are correlated. When personality traits are measured, with tests of approximately the same reliability, the correlations between twins are markedly lower, and there is a slight tendency for correlations of fraternal pairs and of identical pairs to be less different. If the assumption is correct that these personality traits are more affected by environment, then it appears that the operation of environmental forces is such as to produce a pattern of resemblance coefficients unlike that found when intelligence measures are involved. The operation of environmental forces is complex and little understood, but it is apparently such as to *reduce* twin resemblances in personality traits. The difficulty with comparative use of such facts in inference is that different environmental forces operate upon different traits, and they probably operate in different ways for various traits and in different circumstances.

VIII. Environmental Influences

The fundamental question at the present time is not " What differences *can* environment produce? " but rather " With what effect are the *existing environmental forces* actually operating? " Some research workers have been directly concerned with study of the environment of twins, with the aim of answering such a question.

Attempts have been made to ascertain the percentage contributed by heredity and by environment through statistical manipulation of coefficients of twin resemblance. The formulas usually involve the assumption that the nurture influences are approximately equal for fraternal and identical twins. Such an assumption seems untenable to anyone who has had much contact with twins in their own social environments, for it is quite evident that the environments of identical twins are on the average more similar than those of fraternal twins. The identical twins obviously like each other better; they obviously have the same friends more often; they obviously spend more time together; and they are obviously treated by their friends, parents, teachers, and acquaintances as if they were more alike than fraternal twins are. Fraternal twins are more often different physically or mentally; sometimes they are so different that they are differently motivated and do not compete with each other in some spheres. Among the dizygotic twins, boy-girl pairs more often have friendly rivalry; same-sex pairs more often have unfriendly rivalry. Among the identical twins, males more often maintain their equality; females more often show variation in dominance-submission relationships, forms of polarity. Such are the writer's impressions based upon field study of twins. These opinions are supported by the reports of parents of twins, who are frequently aware that fraternal twins are incapable of experiencing and assimilating the same environmental influences.

The fact of such differences in nurture is not generally denied, but is reacted to differently by various workers. Newman and others (84) consider it a relatively minor factor, on the basis of careful treatment of available data. But studies directly planned to investigate such factors are most desirable.

The most important quantitative material bearing on such issues is that provided by Wilson and Jones (133). Wilson (132) has shown that fraternal twins differ more than identical twins in disease histories, and in general health. The differences, such as weight differences, when present in early life, tend to be retained into the school period; hence, they constitute an important basis for cumulative environmental differences based upon hereditary differences. Wilson (132) has further shown that the identical twins spend more time together, have more similar tastes in foods, are more often in the same grade in school, and more often occupy the same room at home; he has shown that these differences are statistically reliable. Jones and Wilson (46) have shown that identical twins are generally considered more similar by their associates; this is interpreted as an environmental factor that must operate toward making fraternal twins more different in personality. Nearly 200 pairs of twins were included in these studies by Wilson and Jones.

From a study of fifteen pairs of monozygotic and eleven pairs of dizygotic twins, von Bracken (4, 5) concludes that the identical twins are more alike in behavior and in interests, and that this constitutes an environmental influence of one upon the other that is more similar for identical twins. He states that identical twins are more congenial, that there is a higher degree of intimacy between them. Stocks (120) has stated that the dizygotic twins are often very different in body build, healthiness, taste, temperament, and the like, so that they naturally experience differences in nurture to a greater degree than monozygotic twins. Geyer (32) has pointed out that the identical twins are more alike even in behavior during sleep. Lehtovaara (55) found in oral questionnaire that identical twins spend more time together, and share the same friends, interests, and ambitions more often than do fraternal twins. Lohmeyer (61, 62) believes that many mistakes are made in the rearing of twins, and that exaggerated attempts to treat both alike are calculated to arouse undesirable competitive attitudes. He states that identical twins sometimes feel underprivileged, and react to the expectation of similarity as a hostile social pressure. The present writer feels that the tendency of identical twins to feel 'cheated out of their individuality' is relatively rare, is largely confined to females, and is a response to an unhealthy sort of expectation on the part of the parents.

To sum up, many of those who have done research on twins have felt that there are nurture influences that are more similar for identical than for fraternal twins, and some workers have made careful quantitative studies of the problem. The insights expressed vary somewhat from author to author, but in view of their possible importance and the fact that these influences have been frequently neglected, it seems desirable to make special note of the suggestions offered. A tendency to take account of this source of stimulation is an emphasized characteristic of recent twin research.

IX. Summary and Interpretation

1. The work of the past ten years has seen increased efficiency in the use of the twin method as an approach to nature-nurture problems. The techniques of diagnosis have become better known, experimental training studies have been undertaken, and increased attention has been paid to the statistical techniques so essential in the treatment and summary of data.

2. The contributions tend to take the form of description, rather than 'explanation,' of the facts of similarity of twins. Knowledge of the resemblances of monozygotic and dizygotic twins is being extended to include traits of character and personality and a more detailed array of mental abilities. The explanation is tending to consist of further description to be used in conjunction with the comparative method in arriving at useful inferences.

3. The whole array of twin-studies seems to suggest, to the writer at least, the futility and artificiality of the idea of untangling nature and nurture influences in the sense of ascertaining the percentage contributions of each in any *general* sense. The view that the idea of percentage contribution can have meaning only for specific mental traits, regarded under rather specific environmental conditions, seems to be implicit in much of the literature. The literature suggests that the hereditary determiners have a more pervasive influence than was heretofore believed; every increase in the body of data brings more evidence of hereditary influence. At the same time, the data indicate that further elucidation of the problems of operation of these hereditary

factors is to be sought in study of the environment. In the description of resemblances, analysis is helpful, but the application of information to understanding of the nature-nurture problem requires synthesis.

4. The most noteworthy study of twins, that of identical twins reared apart, has served as a focal point for attention of students of the nature-nurture problem. The complexities encountered in this study (84) serve to emphasize the statement that progress has been in the form of more perfect understanding of problems, rather than in any dramatic solution of any one problem. The data show one thing clearly, that drastic differences in the educational and social environment are sometimes associated with moderate differences in the IQ's of identical twins reared apart. That only this can be said should bring caution into the writings of the most ardent environmentalists.

References

(1) Ardashnikov, S. N., Lichtenstein, E. A., Martynova, R. P., Soboleva, G. V., and Postnikova, E. N. "The diagnosis of zygosity in twins." *Jour. Heredity,* 27: 1936, 465–468.

(2) Blatz, W. E., Chant, N., Charles, M. W., and Others. *Collected Studies on the Dionne Quintuplets.* (The University of Toronto Press: Toronto, 1937) 238 pp.

(3) Bouterwek, H. "Erhebungen an eineiigen Zwillingspaaren über Erbanlage und Umwelt als Charakterbildner." ["Studies of heredity and environment as factors in the formation of the character of identical twins."] *Z. Menschl. Vererb.- u. KonstLehre,* 20: 1936, 265–275.

(4) Bracken, H. V. "Psychologische Untersuchungen an Zwillingen." *Indus. Psychotechn.,* 10: 1933, 351–352.

(5) Bracken, H. V. "Mutual intimacy in twins: types of social structure in pairs of identical and fraternal twins." *Character and Personality,* 2: 1934, 293–309.

(6) Bradley, I. A. "Manic-depressive psychosis in identical twins." *Amer. Jour. Psychiatry,* 9: 1930, 1061–1064.

(7) Brauns, L. "Studien an Zwillingen im Säuglings und Kleinkindesalter." ["Studies on twins in infancy and preschool age periods."] *Zsch. f. Kinderforsch.,* 43: 1934, 86–129.

(8) Brintle, S. L. "Mental and physical measurements of a set of twelve-year-old quadruplets." *Ped. Sem. and Jour. Genet. Psychol.,* 39: 1931, 91–112.

(9) Brody, D. "Twin resemblances in mechanical ability with reference to the effects of practice on performance." *Child Develop.,* 8: 1937, 207–216.

(10) Byrns, R. "The mental ability of twins." *School and Soc.,* 40: 1934, 671–672.

(11) Carmena, M. "Ist die personliche Affektlage oder 'Nervosität' eine ererbte Eigenschaft?" *Zsch. f. d. ges. Neurol. u. Psychiat.,* 150: 1934, 434–445.

(12) Carmena, M. "Schreibdruck bei Zwillingen." *Zsch. f. d. ges. Neurol. u. Psychiat.,* 152: 1935, 19–24.

(13) CARTER, H. D. "Identical twins reared together." *Jour. Heredity,* 23: 1932, 53–66.
(14) CARTER, H. D. "Twin similarities in occupational interests." *Jour. Educ. Psychol.,* 23: 1932, 641–655.
(15) CARTER, H. D. "Twin-similarities in personality traits." *Ped. Sem. and Jour. Genet. Psychol.,* 43: 1933, 312–321.
(16) CARTER, H. D. "Case studies of mature identical twins." *Ped. Sem. and Jour. Genet. Psychol.,* 44: 1934, 154–174.
(17) CARTER, H. D. "Twin-similarities in emotional traits." *Character and Personality,* 4: 1935–1936, 61–78.
(18) CLARKE, A. E., and REVELL, D. G. "Monozygotic triplets in man." *Jour. Heredity,* 21: 1930, 147–156.
(19) CUMMINS, H. "Dermatoglyphics in twins of known chorionic history, with reference to diagnosis of the twin varieties." *Anat. Rec.,* 46: 1930, 169–198.
(20) CURTIUS, F. "Nachgeburtsbefunde bei Zwillingen und Ähnlichkeitsdiagnose." *Arch. f. Gynak.,* 140: 1930, 361–366.
(21) DAY, E. J. "The development of language in twins. II. The development of twins: Their resemblances and differences." *Child Develop.,* 3: 1932, 298–316.
(22) DIEHL, K., and VON VERSCHUER, O. "Erbuntersuchungen an tuberkulösen Zwillingen." *Beitr. z. Klin. d. Tuberk.,* 75: 1930, 206–215.
(23) ESSEN-MÖLLER, E. "Den moderna twillingforskvingen." ["Modern research on twins."] *Svenska Läkartidn.,* 34: 1937, 900–910.
(24) FAXÉN, N. ["Hypothyroidism in one of twins."] *Acta paediat.,* 17: 1935, 565–573.
(25) FORTUYN, A. B. D. "Modern research on human twins." *Quar. Rev. Biol.,* 7: 1932, 298–306.
(26) FREEMAN, F. N. "Heredity and environment in the light of the study of twins." *Scientific Monthly,* 44: 1937, 13–19.
(27) FREEMAN, W. "Symptomatic epilepsy in one of identical twins; study of epileptic character." *Jour. Neurol. and Psychopath.,* 15: 1935, 210–218.
(28) FRISCHEISEN-KÖHLER, I. "Untersuchungen an Schulzeugnisse von Zwillingen." *Zsch. f. angew. Psychol.,* 37: 1930, 385–416.
(29) FRISCHEISEN-KÖHLER, I. "Zwillinge in den Familien von Hilfsschülern." *Zsch. menschl. Vererb.- u. KonstLehre.,* 20: 1936, 1–8.
(30) FRUMKIN, S. "Mongolian twins." *Proc. Amer. Assoc. Mental Deficiency,* 40: 1935, 253–265.
(31) GESELL, A. "The developmental psychology of twins." [In] *A Handbook of Child Psychology.* C. Murchison, ed. (Clark University Press: Worcester, Mass., 1931) 711 pp. (pp. 158–203).
(32) GEYER, H. "Schlafen und Wachen bei Zwillingen." ["Sleeping and waking in twins."] *Forsch. Fortschr. dtsch. Wiss.,* 13: 1937, 209–291.
(33) HARTMANN, H., and STUMPF, F. "Psychosen bei eineiigen Zwillingen." *Zsch. f. d. ges. Neurol. u. Psychiat.,* 123: 1929–1930, 251–298.
(34) HELLSTEN, H. "Über das Vorkommen der mongoloiden Idiotie, sowie ein Beitrag zur Kasuistik der mongoloiden Zwillinge." *Mschr. Kinderheilk.,* 66: 1936, 400–405.

(35) Herrman, L., and Hogben, L. "The intellectual resemblance of twins." *Proc. Royal Society* (Edinburgh), 53: 1933, 105–129.

(36) Hilgard, J. R. "The effect of early and delayed practice on memory and motor performances studied by the method of co-twin control." *Genet. Psychol. Monog.*, 14: 1933, 493–567.

(37) Hirsch, N. D. M. *Twins: Heredity and Environment.* (Harvard University Press: Cambridge, Mass., 1930) 157 pp.

(38) Holzinger, K. J. "The relative effect of nature and nurture on twin differences." *Jour. Educ. Psychol.*, 20: 1929, 241–248.

(39) Holzinger, K. J. "The statistical evaluation of nature and nurture." *Jour. Amer. Statist. Assoc.*, 30: 1935, 274–280.

(40) Holzinger, K. J. "Reply to special review of 'Twins.'" *Psychol. Bull.*, 35: 1938, 436–444.

(41) Humm, D. G. "Mental disorders in siblings." *Amer. Jour. Psychiat.*, 12: 1932, 239–283.

(42) Hunt, W. A., and Clarke, F. M. "The startle pattern in children and identical twins." *Jour. Exper. Psychol.*, 21: 1937, 359–362.

(43) Jenkins, R. L. "Twin and triplet birth ratios: A further study of the interrelations of frequencies of plural births." *Jour. Heredity*, 20: 1929, 485–494.

(44) Jenkins, R. L. "Dissimilar identical twins; results of brain injury at birth." *Amer. Jour. Orthopsychiat.*, 5: 1935, 39–42.

(45) Jones, H. E. "Twins." *Calif. Monthly*, 24: May, 1930, 24–25.

(46) Jones, H. E., and Wilson, P. T. "Reputation differences in like-sex twins." *Jour. Exper. Educ.*, 1: 1932–1933, 86–91.

(47) Kanaev, I. I. ["Concerning conditioned reflexes in uniovular twins."] *Arkh. biol. Nauk*, 34: 1934, 569.

(48) Kanaev, I. I. "Physiology of the brain in twins." *Character and Personality*, 6: 1937–1938, 177–187.

(49) Kasanin, J. "Case of schizophrenia in only one of identical twins." *Amer. Jour. Psychiat.*, 91: 1934, 21–28.

(50) Kerr, M. "Temperamental differences in twins." *Brit. Jour. Psychol.*, 27: 1936–1937, 51–59.

(51) Komai, T., and Fukuoko, G. "A set of dichorionic identical triplets." *Jour. Heredity*, 22: 1931, 233–243.

(52) Lahdensuu, S. "Mongolismus bei einem der Zwillinge sowie ein Fall mit zwei mongoloiden Kindern in derselben Familie." *Monatsch. Kinderheilk.*, 71: 1937, 14–21.

(53) Lange, J. "Leistungen der Zwillingspathologie für die Psychiatrie." *Allg. Zsch. f. Psychiat.*, 90: 1929, 122–142. *Wien. klin. Woch.*, 42: 1929, 1213–1217; 1247–1250.

(54) Law, F. W. "The refractive error of twins." *Brit. Jour. Ophthal.*, 19: 1935, 99–101.

(55) Lehtovaara, A. *Psychologische Zwillingsuntersuchungen.* (Academiae Scientiarum Fennicae: Helsinki, 1938) 460 pp.

(56) Lereboullet, P. "Les jumeaux: leur physiologie; leur pathologie; l'identité des jumeaux univitellins." ["Disease similarities in identical twins."] *Nourrisson*, 23: 1935, 277–296.

(57) LEVIT, S. G. "Twin investigations in the U. S. S. R." *Character and Personality,* 3: 1935, 188–193.

(58) LEWIS, A. "A case of apparent dissimilarity of monozygotic twins." *Annals of Eugen.,* 7: 1936, 58–64.

(59) LEY, J. "Un cas d'audi-mutité idiopathique (aphasie congénitale) chez des jumeaux monozygotiques." *Encéph.,* 24: 1929, 121–165.

(60) LEY, J., and TORDEUR, G. W. "Alexie et agraphie d'évolution chez des jumeaux monozygotiques." *Jour. belge Neurol. Psychiat.,* 36: 1936, 203–222.

(61) LOHMEYER, G. "Zwillingserziehung." *Int. Zsch. f. Indiv.-Psychol.,* 6: 1931, 461–465.

(62) LOHMEYER, G. "The upbringing of twins." *Internat. Jour. Indiv. Psychol.,* 1: 1935, 113–117.

(63) LURIA, A. R. "The development of mental functions in twins." *Character and Personality,* 5: 1936–1937, 35–47.

(64) LURIA, A. R., and MIRENOVA, A. N. ("Experimental development of constructive activity. Differential training of identical twins.") *Trud. med.-genet. Inst.,* 4: 1936, 487–505.

(65) LURIA, A. R., and MIRENOVA, A. N. ("Investigation of the experimental development of perception by the method of different training of identical twins.") *Neurol. i. Genet.,* 1: 1936, 407–442.

(66) MACARTHUR, J. W., and FORD, N. *A Biological Study of the Dionne Quintuplets: An Identical Set.* (University of Toronto Studies, Child Development Series, No. 11. Toronto, 1937) 49 pp.

(67) McGRAW, M. B. *Growth: A Study of Johnny and Jimmy.* (D. Appleton-Century: New York, 1935) 319 pp.

(68) McNEMAR, Q. "Twin resemblances in motor skills, and the effect of practice thereon." *Ped. Sem. and Jour. Genet. Psychol.,* 42: 1933, 70–99.

(69) McNEMAR, Q. "Special review of Newman, Freeman, and Holzinger's twins: A study of heredity and environment." *Psychol. Bull.,* 35: 1938, 237–249.

(70) MARTINEZ, A., and CIAMPI, L. ["The simultaneous psychosis of the adolescent period."] *Bol. Inst. psiquiat.,* 4: 1933, 91–104.

(71) MERRIMAN, W. E. "Psychoses in identical twins." *Psychiat. Quart.,* 7: 1933, 37–49.

(72) MEUMANN, I. "Testpsychologische Untersuchung an ein-und zweieiigen Zwillingen." *Arch. f. d. ges. Psychol.,* 93: 1935, 42–82.

(73) MIRENOVA, A. N. ["Psychomotor education and the general development of preschool children: Experiments with twin controls."] *Trud. med.-biol. nauchno-issled. Inst. Gorkogo,* 13: 1934, 86–103.

(74) MIRENOVA, A. N. "Psychomotor education and the general development of preschool children: experiments with twin controls." *Ped. Sem. and Jour. Genet. Psychol.,* 46: 1935, 433–454.

(75) MIRENOVA, A. N., and KOLBANOVSKI, V. N. ["A comparative evaluation of methods for the development of combinative functions in preschool children."] *Trud. med.-nauchno-issled. Inst. Gorkogo,* 3: 1934, 104–118.

(76) NEWMAN, H. H. "Studies of human twins: I. Methods of diagnosing monozygotic and dizygotic twins." *Biol. Bull.,* 55: 1928, 283–297.

(77) NEWMAN, H. H. "Studies of human twins: II. Asymmetry reversal or mirror image in identical twins." *Biol. Bull.*, 55: 1928, 298–315.

(78) NEWMAN, H. H. "Identical twins: The differences between those reared apart." *Eug. Rev.*, 22: 1930–1931, 29–34.

(79) NEWMAN, H. H. "Palmar dermatoglyphics of twins." *Amer. Jour. Physical Anthropol.*, 14: 1930, 331–378.

(80) NEWMAN, H. H. "Palm-print patterns in twins; on the use of dermatoglyphics as an aid in the diagnosis of monozygotic and dizygotic twins." *Jour. Heredity*, 22: 1931, 41–49.

(81) NEWMAN, H. H. "Identical twins." *Scientific Monthly*, 34: 1932, 169–171.

(82) NEWMAN, H. H. "Mental and physical traits of identical twins reared apart: Case IV. Twins Mary and Mabel, and review of the first four cases studied." *Jour. Heredity*, 23: 1932, 3–18.

(83) NEWMAN, H. H. "The effects of hereditary and environmental differences upon human personality as revealed by studies of twins." *Amer. Naturalist*, 67: 1933, 193–205.

(84) NEWMAN, H. H., FREEMAN, F. N., and HOLZINGER, K. J. *Twins: A Study of Heredity and Environment.* (University of Chicago Press: Chicago, 1937) 369 pp.

(85) OREL, H. "Längen- und Massenwachstum von Zwillingen. Kleine Beiträge zur Vererbungswissenschaft." ["Growth in length and weight in twins. VII. The knowledge of heredity."] *Zsch. f. Kinderh.*, 48: 1929, 1–14.

(86) ORGLER, H. "Identische, getrennt aufgezogene Zwillinge." *Int. Zsch. Indiv.-Psychol.*, 13: 1935, 35–36.

(87) PARKER, P. A. "Variations in the mental development of twins." *Proc. and Addr. Amer. Assoc. Study Feebleminded*, 56: 1932, 213–221.

(88) PINCHERLE, B. "Diverso comportamento di due gemelli monocorii di fronte alla sifilide materna." ["Opposite behavior of monochorionic twins with reference to maternal syphilis."] *Pediatria*, 42: 1934, 1207–1220.

(89) POPENOE, P. "Twins and criminals." *Jour. Heredity*, 27: 1936, 388–390.

(90) REED, G. E. "Uniovular twins; schizophrenia and tuberculosis." *Canadian Med. Assoc. Jour.*, 32: 1935, 180–182.

(91) REICHLE, H. A. "The diagnosis of monovular twinning." *Biol. Bull.*, 56: 1929, 313–326.

(92) REICHLE, H. A. "The diagnosis of the type of twinning. I. Dermatoglyphics." *Biol. Bull.*, 56: 1929, 164–176.

(93) RICHARDSON, S. K. "The correlation of intelligence quotients of siblings at the same chronological age levels." *Jour. Juv. Res.*, 20: 1936, 186–198.

(94) RIFE, D. C. "Genetic studies of monozygotic twins." *Jour. Heredity*, 24: 1933, 339–345, 407–414, 443–446.

(95) RIFE, D. C. "Twin differences in intelligence." *Jour. Educ. Psychol.*, 26: 1935, 709–712.

(96) RIFE, D. C. "Contributions of the 1937 National Twins' Convention to research." *Jour. Heredity*, 29: 1938, 83–90.

(97) ROSANOFF, A. J. "A study of mental disorders in twins." *Eug. News*, 17: 1932, 37–39.

(98) ROSANOFF, A. J. "Some clinical manifestations of traumatic decerebration." *Psychiat. Quart.*, 9: 1935, 116–128.

(99) ROSANOFF, A. J., and HANDY, L. M. "The etiology of mongolism with special reference to its occurrence in twins." *Amer. Jour. Diseases of Children,* 48: 1934, 764–779.

(100) ROSANOFF, A. J., HANDY, L. M. and PLESSET, I. R. "The etiology of manic-depressive syndromes with special reference to their occurrence in twins." *Amer. Jour. Psychiat.*, 91: 1935, 725–762.

(101) ROSANOFF, A. J., HANDY, L. M., and PLESSET, I. R. "The etiology of mental deficiency with special reference to its occurrence in twins: A chapter in the genetic history of human intelligence." *Psychol. Monog.*, 48: 1937, No. 4, 1–137.

(102) ROSANOFF, A. J., HANDY, L. M., and ROSANOFF, I. A. "Criminality and delinquency in twins." *Jour. Crim. Law and Crimin.*, 24: 1934, 923–934.

(103) ROSANOFF, I. A., and ROSANOFF, A. J. "A study of mental disorders in twins." *Jour. Juv. Res.*, 15: 1931, 268–270.

(104) RUSSELL, P. M. G. "Mongolism in twins." *Lancet,* 224: 1933, 802–803.

(105) SANDERS, J. "Similarities in triplets." *Jour. Heredity,* 23: 1932, 225–234.

(106) SANDERS, J. "Twin research in Holland." *Jour. Heredity,* 23: 1932, 499–502.

(107) SANDERS, J. "Zwillinge mit Epilepsie." *Genetica,* 15: 1933, 67–88.

(108) SAUDEK, R. "Identical twins reared apart: A comparative study of various tests of their intellectual, emotional, and social attitudes." *Character and Personality,* 2: 1933, 22–40.

(109) SAUDEK, R. "A British pair of identical twins reared apart." *Character and Personality,* 3: 1934, 17–39.

(110) SCHIFF, F., and VON VERSCHUER, O. "Serologische Untersuchungen an Zwillingen." ["Serologic investigation of twins."] *Klin. Wochnschr.,* 10: 1931, 723–726.

(111) SCHILLER, M. "Zwillingsprobleme, dargestellt auf Grund von Untersuchungen an Stuttgarter Zwillingen." *Zsch. menschl. Vererb.- u. Konst. Lehre,* 20: 1936, 284–337.

(112) SEEMAN, E., and SAUDEK, R. "The self-expression of identical twins in handwriting and drawing." *Character and Personality,* 1: 1932, 91–128.

(113) SEEMAN, E., and SAUDEK, R. "The handwriting of identical twins reared apart." *Character and Personality,* 1: 1933, 268–285.

(114) SHATTUCK, R. H. "Mongolism in one twin." *Eug. News,* 14: 1929, 134–135.

(115) SMITH, J. C. "Das Ursachenverhältnis des Schwachsinns beleuchtet durch Untersuchungen von Zwillingen." *Zsch. f. d. ges. Neurol. u. Psychiat.,* 125: 1930, 678–692.

(116) SMITH, J. C. "Psychiatrische Zwillingskasuistik." *Acta psychiat. et neurol.,* 6: 1931, 79–105.

(117) SOBOLEVA, G. V. ["The study of twins in Moscow."] [*Russian Jour. Eug.*] 4: 1926, 3–22.

(118) SONTAG, L. W., and NELSON, V. L. "A study of identical triplets: Part I. Comparison of the physical and mental traits of a set of monozygotic dichorionic triplets." *Jour. Heredity,* 24: 1933, 474–480.

(119) STERN, W., and KATZENSTEIN, B. "Verschiedene Glaubwürdigkeit siebenjähriger Zwillingsschwestern." *Zsch. f. angew. Psychol.,* 40: 1931, 231–238.

(120) STOCKS, P. "A biometric investigation of twins and their brothers and sisters." *Annals of Eugenics,* 4: 1930, 49–108.

(121) STRAYER, L. C. "Language and growth: The relative efficacy of early and deferred vocabulary training, studied by the method of co-twin control." *Genet. Psychol. Monog.,* 8: 1930, No. 3, 209–319.

(122) STUMP, N. F. "Similarity of identical twins in learning a pencil maze." *Jour. Heredity,* 27: 1936, 352–354.

(123) VON VERSCHUER, O. "Die Variabilität des menschlichen Korpers an Hand Wachstumsstudien an ein- und zweieiigen Zwillingen." *Zsch. Indukt. Abstamm.- u. Vererbungsl., Supplement* 2: 1928, 1508–1516.

(124) VON VERSCHUER, O. "Demonstrationen zur Zwillingsanthropologie." *Verhandl. d. Gesellsch. f. phys. Anthropol.,* 4: 1930, 67–74.

(125) VON VERSCHUER, O. "Erbpsychologische Untersuchungen an Zwillingen." *Zsch. f. induktive Abstamm.- u. Vererbungslehre,* 54: 1930.

(126) VOÛTE, P. A. ["Research on twins."] *Maandschr. Kindergeneesk.,* 5: 1936, 202.

(127) WARING, J. I. "Mongolism in one of twins." *Amer. Jour. Dis. Child.,* 41: 1931, 351–353.

(128) WATSON, E. M. "Diabetes mellitus in twins." *Canadian Med. Assoc. Jour.,* 31: 1934, 61–63.

(129) WHITE, K. B. "A clinical study of twenty-six pairs of twins." *Psychol. Clin.,* 21: 1932, 243–252.

(130) WIGERS, F. "Ein eineiiges, bezüglich Schizophrenie diskordantes Zwillingspaar." *Acta psychiat. et neurol.,* 9: 1934, 541–556.

(131) WILSON, J. A. "Twins with eye defects. Ametropia and strabismus." *Brit. J. Ophthal.,* 16: 1932, 421–423.

(132) WILSON, P. T. "A study of twins with special reference to heredity as a factor determining differences in environment." *Human Biol.,* 6: 1934, 324–354.

(133) WILSON, P. T., and JONES, H. E. "A study of like-sexed twins. I. The vital statistics and familial data of the sample." *Human Biol.,* 3: 1931, 107–132.

CHAPTER IX

INTELLIGENCE AS RELATED TO RACE

LETA S. HOLLINGWORTH
Professor of Education
Teachers College, Columbia University
New York, New York
and
PAUL WITTY
Professor of Education
Northwestern University
Evanston, Illinois

I

THE PROBLEM OF COMPARING RACES

LETA S. HOLLINGWORTH

Any complete summary of knowledge about intelligence should give attention to the question as to whether intelligence varies with the biological variations involved in ' race.' The fact seems to be that not one but several sciences deal technically with race and are mutually interdependent for progress in their several fields of specialization. It is necessary to consider here some of these interrelations.

I. THE POSITION OF ANTHROPOLOGY

At the 1938 annual meeting of the American Anthropological Association (13), a resolution was unanimously adopted, in part as follows:

> (1) Race involves the inheritance of similar physical variations by large groups of mankind, but its psychological and cultural connotations, if they exist, have not been ascertained by science. . . .
>
> (3) Anthropology provides no scientific basis for discrimination against any people on the ground of racial inferiority, religious affiliation, or linguistic heritage. (p. 30)

The position of American anthropology is thus clearly stated. It is that psychologists have not yet ascertained whether differences in mental ability or in mental pattern exist among the races of men. More specifically, the position would be that it is at present wholly unknown whether there are differences in *intelligence* as related to ethnic species.

The passing of such a resolution by a scientific body is unusual and suggests that the subject upon which pronouncement is made must be deemed to have some extraordinary significance. In this attitude, that the question of race differences in mentality is of large significance, the writers of this chapter, and of this Yearbook as a whole, would concur.

It would also probably be agreed by psychologists that the psychological connotations of race have not been ascertained, and that a scientific basis is lacking for discrimination against any people on the ground of racial inferiority. And, furthermore, it should be added, as the anthropologists failed to add, that neither does science provide a basis for not discriminating against any people on the ground of racial inferiority. In fact, science at present does not cover this field at all, though tentative excursions into it have been attempted.

In the absence of guidance by experts, people dealing with the practical problems called 'racial' can do no better than follow their feelings, and judge by 'common sense.' Emotion and common sense being but inadequate guides, it is important for science to collect data in this area.

II. The Present Status of the Concept of Race

In thinking about the comparative intelligence of races, it is necessary first to define 'a race.' Some psychologists take the view that this is so difficult a matter as almost to preclude investigation of comparative racial intelligence (1).

> It is apparent that no one criterion of race can yield a satisfactory classification. Nor can clear-cut group distinctions be made with a combination of such criteria. It should be borne in mind that at best any racial classification is approximate. No sharp line of demarcation can be established between groups, nor can every individual be unequivocally assigned to one particular group. . . . An additional difficulty in the way of racial classification is introduced by the extensive amount of race mixture, which has been going on for countless generations. (pp. 461–465) [1]

However imperfect the definitions of race may be, as furnished to us by anthropology, it nevertheless seems quite possible for distinguishable and namable congeries of persons to be defined as belonging to

[1] Quoted by permission of The Macmillan Company, publisher.

census groups, and to be studied as regards comparative intelligence by appropriate methods. This is, of course, a less ambitious enterprise than the determination of the intelligence of whole races, but it has at least the merit of feasibility, and it is valuable for the purposes of school and society.

A 'census group,' in the terms just described, would be the Cherokees of the Five Civilized Tribes of Oklahoma, or school children of Welsh descent in Wilkes-Barre, Pennsylvania. These groups would be defined in part on the basis of anthropology, in part on the basis of ecology, and in part on the basis of political science. It is upon groups thus distinguishable and namable that a large number of studies have been made in the United States in relation to intelligence, with the result that many differences have been found among such comparative groups (5), as well as many similarities. This is, however, essentially different from studying *races*.

For purposes of classification by race, preliminary to investigation of comparative racial intelligence, some *combination* of physical (or of mental and physical) traits may finally be evolved by anthropometrists. The really crucial structures for inclusion in this defining combination will ultimately appear in the course of long-continued research. Is there, for instance, a typical Negro foot? Is there a typical Chinese eye?[1] Is there a typical Sioux nose? Each of these structures is a highly complex affair. No single variable, such as 'length' or 'width,' would serve for effective study of the problem. What is needed is reference to a *configuration* involving contour, color, and other elements that at present completely elude morphometry. Concepts like 'brachycephalic' and 'dolicocephalic' are first steps in the right direction.

The common man in South Dakota knows 'an Injun' when he sees one. Judges chosen at random from persons dwelling within 50 miles of Rosebud Reservation would agree, almost to a man, as to whether an individual presented for inspection is or is not 'an Injun.' It is the *configuration* that tells the story, but configuration is as yet too difficult for anthropometry to grasp analytically. Only the human eye grasps it, assisted by the rest of the sensory equipment.

Take, for instance, the great difficulty of measuring the hand. This structure is so complex that no apparatus has been devised for recording it as a whole. The nearest approach to an attempt in this direction is the 'Glove Fit,' constructed at Colgate University for use in making and selling gloves,

[1] 'The smooth eye-lid' was mentioned several times by Chinese engaged in judging the comparative beauty of adolescent faces (12).

and in measuring glove-sizes. Until anthropometry and anthropology pass into a stage of more adequate insight, the ambitious project of studying *race* differences in intelligence will be in abeyance.

Before passing from this subject, we may remark that there is no reason why the definition of race should be confined ultimately to physical variations. If by research it should be determined that there are mental configurations typical of ethnic stocks, these, too, might be included in the definition of a race. Researches like that of Hunter (10), using the Rorschach method of studying personality in the comparison of census groups, are steps toward this goal. At present, little or nothing of such patterns is known to science, but the common man alleges that he recognizes them, and he is influenced in his practical behavior by them. " Once there was an Irishman," for instance. It is at present wholly unknown as to whether science will ever verify or not the psychological pattern called ' Irish,' but it would be worth while to see whether the common man is suffering from an illusion. It is possible that if enough elements could be included in a configuration that science might handle, little or no overlapping would be found among races in the distribution of *pattern;* so that ultimately the anthropologist and the psychologist would become at least as able as the common man is to relate an individual to his race or race-mixture.

Because of the largeness and the controversial aspects of the subject, it has been decided for purposes of this Yearbook merely to comment here upon racial characteristics as a problem, rather than to review the literature of research; to offer a concrete sample of current discussion of one census group (popularly called a ' race ') by a psychologist; to circumscribe some areas of ignorance that invite research upon comparative intelligence; and finally to recommend that in the proper course of events a whole Yearbook in this series be devoted to the subject of racial characteristics, especially in regard to comparative mentality, as the matter bears upon the interests of school and society. There will be no attempt to exhaust the subject here or to synthesize current knowledge about it.

III. The Concept of Census Groups

Abandoning the idea that mental characteristics or qualities as inherent in race have been investigated, it still remains true that the comparative intelligence of what we have called ' census groups ' is open to study. Census groups may be chosen on an ethnic basis, as having in common certain physical variations that cause them to be officially classified together. The Japanese, in California, would exemplify such a group. Or such groups may be chosen as claiming, and being verified as having, a common origin of the surname, as, for instance, the French-Canadians. Or the census group may be chosen for purposes of comparative psychological study on the basis of having a European origin in common, as, for instance, Americans of Spanish descent.

That such groups are distinguishable and namable cannot be disputed, and that, so classified, persons can be studied and compared as regards the factor of intelligence is obvious. It has already been done to a considerable extent, and the data resulting are verifiable and are of practical interest for educators.

If other methods of selection should fail or prove unsatisfactory, the methods of judgment employed in the psychological laboratory to objectify intangibles might be used, and groups might be chosen by the judgments of common men, using the method of paired comparisons or one of the other standard methods of judgment. Thus the same methods might be used to determine who belong to a given ethnic group as have been used to determine whether or not twins are 'identical.' The difficulties of selecting census groups are by no means insurmountable.

In the discussion immediately following, we have an example of a definite census group, the American Negro. The discussion of this one group exemplifies the kind of useful research that is feasible in the comparison of ethnic congeries in the United States today. It points the way to the kinds of studies needed from psychologists, while anthropology, ecology, and biology are engaging in their attempt to furnish a definition of *race* that will finally pave the way for a study of *racial* comparisons.

II

RESEARCH UPON THE AMERICAN NEGRO

PAUL WITTY

I. STUDIES OF THE INTELLIGENCE OF NEGROES AND NEGRO-WHITE MIXTURES

Rates of mental development have been assigned to various 'races' in America. The procedure typically followed has been to test contiguous samples of each 'racial' group, calculate average scores, and make comparisons after consulting age equivalents or norms. Thus,

C. C. Brigham (3) and others analyzed the scores made by foreign-born recruits in the United States Army during the World War; the recent immigrants made relatively lower average scores than the older immigrant stocks. The conclusion was drawn that the average intelligence of the Italians and the Russians, for example, was low, while that of the English, the Germans, and the native-born Americans was high. This conclusion has been disproved, and assertions have been qualified or retracted (2). But doctrines of racial superiority and inferiority persist. They have had unusually wide and far-reaching application in the case of the American Negro.

Negroes in America have been studied much more extensively than other 'racial' groups; they invariably make lower average scores than whites. Moreover, groups of Negroes possessing allegedly larger amounts of white blood usually make higher average scores than those of (suspected) pure Negroid ancestry.

Results from extensive and thorough research have demonstrated, of course, that there are *differences between the races, and in subgroups within each race — not that there are true racial differences in innate or inherited intelligence* (11).

There are very few published studies dealing with the relationship of Negro-white ancestry to intelligence. Investigators have usually separated the Negro subjects into groups on the basis of skin color and compared average intelligence-test scores. This method is of doubtful validity, because it has been demonstrated that the racial composition of the individual Negro cannot be determined accurately by estimates of pigmentation (8, 9).

Ferguson's (4) investigations are well known in this field. In one study he classified Negro school children in four groups: "pure," "three-fourths Negro," "mulatto," and "quadroon," on the basis of "skin color, hair texture, and general facial and cranial conformation, the main emphasis being upon skin color." The magnitude of the average intelligence-test scores was inversely related to the amount of pigmentation in the groups. Similarly, Young (18) found "a noticeable decrease of intelligence as we go from the white children to light Negroes and then to dark Negroes." A study by Strong (14) is of interest, since results differ from those usually reported. Negro children were subdivided into: the dark, the medium, and the light. Paradoxically, the dark children made higher scores than the light!

These investigations have serious defects and limitations. Herskovits' anthropometric measures of Howard University males, reported in 1926 (8), illustrate a rather thorough approach to the problem of the relationship of Negro-white mixture and mental-test ability. Measures of differentiating physical traits, such as width of nostril, lip thickness, and pigmentation, were correlated with scores on the Thorndike College Entrance Examination. The coefficients of correlation, based on 115 cases, varied from +.014 to −.198. Herskovits concludes:

> . . . The relationship between test scores and physical traits denoting greater or less amounts of Negro blood is so tenuous as to be of no value in drawing conclusions as to the comparative native ability or relative intelligence of the Negro when compared to the White. (p. 41)

The uncritical student has made, however, sweeping generalizations concerning the 'lack of educability' and the general constitutional inferiority of Negro children. One leaves the literature with the impression that the Negro child constitutes hopeless school material. In addition, one might almost conclude that gifted Negro children are so rarely found in the public school that search for them would prove unprofitable.

II. The Study by Jenkins of Negro Children of Superior Intelligence

Several recent studies have demonstrated the falsity of these assumptions. In what follows, salient data from an investigation by Jenkins will be set forth (11). This investigation was carried out under the direction of the writer, who participated in making and reporting certain phases of it.

1. Selection of Children

The children were identified in a systematic search for superior Negro children in Grades III to VIII of seven Chicago public schools in which approximately 8,000 Negro children were enrolled.[1] Class-

[1] Jenkins states that an area for such a study must conform to certain specifications: first, a large number of Negro children must be available for testing; second, a large number of the parents must be in the higher occupational levels; and third, sufficient opportunity for educational and cultural development must be available. Chicago met these conditions, and the schools of that city were selected on the basis of socio-economic zones taken from E. F. Frazier's, *The Negro Family in Chicago*. Of the seven schools selected, four were from the best zones; two, from the average zones; one, from an inferior zone.

room teachers nominated: (*a*) the child thought most intelligent, (*b*) the child doing the best classroom work, and (*c*) children one or more half-years underage for their grades.

Five hundred and thirty-nine children, approximately 6.5 percent of their school population, were nominated. The McCall Multi-Mental Scale was given to 512 of these nominees and the Stanford-Binet examination to every child who had been credited with an IQ of 120 or more on the McCall Scale. One hundred three children possessing a Stanford-Binet IQ of 120 or above were thus located.

2. Tests and Instruments Used

The following tests and instruments were administered:

The New Stanford Achievement Test, Advanced Examination, Form W
The Personal Index (a battery of four character and personality tests)
The Pupil Report (a questionnaire concerning pupil interests and activities)
The Sims Score Card for Socio-Economic Status

In addition, each child was rated by one teacher on traits like leadership and originality; also, information relative to the hereditary and environmental background, developmental history, school progress, interests, and aptitudes was secured by interviews with parents.

3. Incidence of High IQ Children in the Total School Population

The percentages of the 8,400 children who scored at or above IQ 120 on the McCall and the Stanford-Binet tests follow:

At or above an IQ of	*McCall Scale*	*Stanford-Binet*
140	0.40	0.33
130	1.11	0.66
120	2.11	1.23

The number of gifted children (IQ 140 or above) discovered in the several schools varied. In one school not a single gifted child was identified, but in four other schools the percentages of gifted children (IQ 140 or above) were .83, .80, .57, and .41, respectively.

4. Distribution in Terms of Age and Grade Levels

No grade contributed a noticeably disproportionate number of children, although there were 24 children in Grade III; 19 in Grade VIII; and 12 in Grade VI. The mean grade level was 5.3.

5. The Range in IQ's

The Stanford-Binet IQ's range from 120 to 200, with a mean of 134.2. This distribution follows:

IQ	*Frequency*
200 to 209	1
190 to 199	0
180 to 189	0
170 to 179	0
160 to 169	1
150 to 159	9
140 to 149	18
130 to 139	29
120 to 129	45
Total	103

6. Sex Differences

The mean IQ of the boys was 134.6 (S.D. = 10.8); that of the girls, 133.9 (S.D. = 13.0). Girls displayed greater variability. The highest IQ, 200, was earned by a girl.

The group contained 72 girls and 31 boys — a ratio, girls to boys, of 233:100. A preponderance of girls was found in all but one of the schools. The ratio of girls to boys in the population of the schools was 114:100; the ratio of girls to boys for pupils nominated was 164:100; and for pupils who earned McCall IQ's of 120 or above, the ratio of girls to boys was 231:100.

7. Hereditary and Environmental Background

The parents constituted a well-educated group. The median father had 13.9 years of schooling and the median mother 12.8 years.

Of the fathers, 33.4 percent were classified in Group V (the highest classification) of the Taussig Scale, and 36.5 percent were in Group IV. Only 6.3 percent appeared in the lowest classification (Group I).

In socio-economic status the parents ranked high when the socio-economic level was estimated by such factors as occupation, income, and ownership of certain commodities, as well as when it was quantitatively appraised by means of the Sims Score Card for Socio-Economic Status. The mean score on the Sims Score Card was 18.7 (S.D. = 6.1).

8. Developmental History

Negro children of superior intelligence appear to conform to the general pattern of children of superior intelligence.

The average family contained 3.0 children; half of the group were first-born children and 26.2 percent were 'only' children. The mean age of walking (several steps alone) was 11.9 months (S.D. = 2.5), and the mean age of talking (short sentences) was 12.9 months (S.D. = 3.3). The average child was in height 2.1 inches and in weight 5.3 points above the norms for his chronological age.

9. School Progress

This group of superior children displayed eight times the amount of underageness found in the total school population from which it was selected. Of the group 67 percent had skipped at least one half-grade; the mean number of half-grades skipped was 1.2. The progress quotient was 119.

10. Educational Achievement

The average child had mastered educational subject matter (as measured by the test) 2.7 grades above the norm for his chronological age and 1.1 years above the norm for his grade placement.

The highest subject quotient was in reading (Mdn. = 135.8, S.D. = 19.9), and the lowest was in arithmetic (Mdn. = 122.0, S.D. = 12.8). The arithmetic quotient was lower, relatively, than that of groups of superior children studied by other investigators.

In general, the achievement of the boys was superior to that of the girls; the mean score of the boys exceeded that of the girls in all subjects except spelling and language usage. In only one test, history and civics, was the difference statistically significant. The mean EQ of the boys was 129.8 (S.D. = 11.6); and that of the girls, 125.9 (S.D. = 9.8).

11. Character and Personality Measures

On each of nine traits, teachers rated superior Negro children above the average. And on a battery of character and personality tests, the subjects made average scores on all tests superior to those of a control group.

12. Racial Composition

In racial composition, based on genealogical information furnished by parents, the subjects were found to represent a typical cross section of the general Negro population. Of the children 22 percent reported no record of white ancestry; 46.1 percent have more Negro ancestry than white; and 15.9 percent have more white ancestry than Negro. Thirty-five percent of the group reported Indian ancestry.

13. Implications of the Study

In view of the relative frequency of gifted children (IQ 140 and above) in several of the schools included in this survey, it is singular that the gifted Negro child has heretofore been considered an anomaly. Of significance to American educators is Jenkins' evidence that we may discover large numbers of children of superior intelligence who are unrecognized and who are being denied the type of educational experiences necessary to their fullest development.

Sufficient facts have been set forth in his study to indicate that superior Negro children exhibit the same characteristics that typify other 'racial' groups of superior children. In home background, developmental history, school progress, educational achievement, and social and personal traits, superior Negro children resemble other groups of American children who are superior in test-intelligence.

This study has far-reaching implications for research in the field of 'race' differences. It may be seen at once that there are a number of very significant investigations bearing on the topic of race intelligence that might be profitably undertaken. The following are suggested as especially worthy:

1. A genetic study of a large number of gifted Negroes from representative districts in the United States. The methodology in identification and follow-up employed by Terman in *Genetic Studies of Genius* might be used fruitfully.
2. Comparative studies of the mental ability of Negro children living under widely varied socio-economic conditions.
3. Studies of the mental development and growth of Negro children during infancy and early childhood. Comparisons of Negro children in 'good' and 'poor' homes might yield significant results.
4. Developmental studies of the interests, activities, and attitudes of Negro children in different IQ groups.
5. Developmental studies of the Negro child's emotional growth in relation to his mental development.
6. Studies of the growth and development of Negro children in foster homes.
7. Investigations using improved anthropometrical and psychological measures to disclose the relationship of intelligence to 'racial' composition. Studies of this character should be made with very young children.

8. Studies of the Negro child's special problems and developmental trends during adolescence.

9. Studies of the amount and character of the mental growth of Negro school children in different school 'atmospheres.'

10. Experiments in preschool education to ascertain the effect of this experience upon IQ.

11. Studies of the effects upon IQ of educational deprivations of differing degrees of severity at different age levels.

12. Studies seeking to develop more adequate techniques to measure the intelligence and the educational attainment of the Negro child.

13. Evaluation studies of the growth and adjustment of Negro children.

14. Studies seeking to provide curricula in terms of the special problems and needs of Negro children.

References

(1) Anastasi, A. *Differential Psychology: Individual and Group Differences in Behavior.* (Macmillan: New York, 1937) 615 pp.

(2) Brigham, C. C. "Intelligence tests of immigrant groups." *Psychol. Rev.*, 37: 1930, 158–165.

(3) Brigham, C. C. *A Study of American Intelligence.* (Princeton University Press: Princeton, 1923) 210 pp.

(4) Ferguson, G. O. "Psychology of the Negro: An experimental study." *Arch. Psychol.*, 5, No. 36: 1916, 1–138.

(5) Garth, T. R. "Racial minds." *Psyche,* 8, No. 3: 1928, 63–70.

(6) Garth, T. R. *Race Psychology.* (Whittlesey House, McGraw-Hill: New York, 1931) 260 pp.

(7) Goodenough, F. L. "Racial differences in the intelligence of school children." *Jour. Exper. Psychol.*, 9: 1926, 388–397.

(8) Herskovits, M. J. "On the relation between Negro-white mixture and standing in intelligence tests." *Ped. Sem. and Jour. Genet. Psychol.*, 33: 1926, 30–42.

(9) Herskovits, M. J. *Anthropometry of the American Negro.* (Columbia University Press: New York, 1930) 283 pp.

(10) Hunter, Mary. "A study of the Rorschach 'Erlebniss-Typus' of comparable white and Negro subjects." Honolulu, Hawaii, 1939. 138 pp. (Privately printed.)

(11) Jenkins, M. D. "A socio-psychological study of Negro children of superior intelligence." (Northwestern University, unpublished doctor's dissertation: Evanston, Ill., 1935)

(12) Madden, R., and Hollingworth, L. S. "How one race judges another for physical attractiveness." *Jour. Soc. Psychol.*, 3: 1932, 463–469.

(13) The New York Meeting of the American Anthropological Association. *Science* (New Series), 89: 1939, 29–30.

(14) Strong, A. C. "Three hundred fifty white and colored children measured by the Binet-Simon measuring scale of intelligence: A comparative study." *Ped. Sem.*, 20: 1913, 485–515.

(15) WITTY, P. A., and JENKINS, M. D. "Case of 'B'—a gifted Negro child." *Jour. Soc. Psychol.,* 6: 1935, 117–124.

(16) WITTY, P. A., and JENKINS, M. D. "The educational achievement of gifted Negro children." *Jour. Educ. Psychol.,* 25: 1934, 585–597.

(17) WITTY, P. A., and JENKINS, M. D. "Intra-race testing and Negro intelligence." *Jour Psychol.,* 1: 1936, 179–192.

(18) YOUNG, P. C. "Intelligence and suggestibility in whites and Negroes." *Jour. Comp. Psychol.,* 9: 1929, 339–359.

CHAPTER X

PERSONALITY AND ADJUSTMENT AS DETERMINERS AND CORRELATES OF INTELLIGENCE

LETA S. HOLLINGWORTH
Professor of Education
Teachers College, Columbia University
New York, New York
with the coöperation of
IRVING LORGE, FREDERICK H. LUND, S. D. SHIRLEY SPRAGG,
HAROLD M. WILLIAMS, and KURT LEWIN

I

INTELLIGENCE AS AN ELEMENT IN PERSONALITY

LETA S. HOLLINGWORTH

I. THE DEFINITION OF PERSONALITY

The attempt to gain a generally acceptable concept of personality is conspicuously in progress in current thinking. Outstanding are the recent contributions of Allport (3), and of Briggs (8), and the record of deliberations of the colloquium conducted under the auspices of the American Psychiatric Association (4).

'Personality' has been variously defined, but throughout these definitions a certain amount of uniformity is noticeable. In the first place, there is the idea of *a totality* of elements, and in the second place, there is the idea of the primary significance of the interaction of this totality in relationships between the person *and other persons.* Our question here is: How does intelligence characteristically affect the pattern of this totality, of which it is one element? And how, if at all, is intelligence affected by the totality?

II. ISOLATION AS A FUNCTION OF EXTREME INTELLIGENCE

The whole problem as to how intelligence determines personality has been approached to some extent through the study of intellectual deviates, references to which are to be found in this Yearbook. A noteworthy observation is that persons who deviate widely from the mean of human intelligence tend to become 'isolates.' This fact holds for both plus and minus deviates, and tends strongly to become established, regardless of what the other elements of personality may be. Intelligence far above the average tends to isolate the individual psychologically (35, 68), although he typically remains situated in society. This tendency is especially in evidence during childhood, while the person is under the control of other persons, and while at the same time habits of feeling and of attitude are being strongly formed. On the other hand, intelligence far below the average tends to isolate the individual not only psychologically, but also socially by segregation in institutions (21).

It is a nice problem to determine the amount of deviation that marks the point at which isolation begins. The hypothesis is here offered that at approximately —3 PE and at approximately +6 PE, in terms of IQ (Stanford-Binet) the distance from 'common sense' begins to be felt in a way that produces isolation from contemporaries, and from the common life of other persons.

Why should isolation begin sooner on the minus side than on the plus side? The answer is to be found in the nature of school and society. Deviation in the direction of inadequacy of comprehension begins to take effect when it is perceived *by other persons,* whereas deviation in the direction of superior adequacy does not begin to take effect on personality until it is felt *by the individual himself.* The interaction between what is felt by others and what is felt by the individual himself is, of course, also effective. The difference just described is really one of emphasis only, and not a clean-cut difference of perception by others versus perception by the self. This difference of emphasis is, however, far-reaching in its implications, which cannot be fully brought out here.

III. The Consequences of Isolation

It follows, from what has been said, that intelligence within a fairly wide range around the mean, and hence including most individuals, finds its function merely as a workable component of personality, with no more importance than many other components, and with no consequent eccentric pattern of the whole. Intelligence in the personality of the average man is 'good for' those functions that are ascribed to it in the discussion in this Yearbook, and elsewhere. In the totality of elements, it enables a person to interact with other persons in such a way that mutual understanding with its natural sequelae results. In so far as this is achieved, the personality is said to be 'normal' and 'adjusted.'

But when the magnitude of the intelligence quotient is greatly reduced or increased beyond that of the vast majority who determine 'common sense,' a lack of 'common sense' arises. By this is meant that the individual thinks and acts differently from the majority. What are some of the probable consequences to personality as a whole, from the resulting isolation, in the case of persons of extremely superior intelligence?

IV. Personality Development in Consequence of Being Highly Intelligent

Persons who are as far away from average men on the plus side as idiots are on the minus side have a number of special problems of personality development, of which a few may be mentioned here. *Negativism* toward authority tends to develop, through perception of the fact that authority is often irrational and erroneous in its operation. *Intolerance* may arise from observing over a long period of time the relatively inept thinking of others. The failure to master the art of suffering fools gladly may result in misanthropy. The perception of the comparative lack of intelligence in the majority of associates may also lead to habits of *chicanery,* which may become fixed as a way of life by the necessity to invent means of following out esoteric interests in a world that is on the whole unadapted to them.

For the very high deviate, *the problem of evil* comes to trouble the mind at an early age, when emotional development and life experience are still greatly limited as compared with the average. This involves emotional conflicts that ordinarily wait on greater maturity of the organism as a whole, and, that therefore tend more strongly to disruption in the child of great intelligence.

Being isolated in interests, vocabulary, and problems from his contemporaries in his immaturity, the highly intelligent child turns to solitary pursuits and to the companionship of older persons. The play of children who test about 170 IQ (Stanford-Binet) is typically solitary. This was noticed long ago by Yoder (69), and has been confirmed by modern investigators (67). The consequences of habituation to older persons, instead of to contemporaries, as companions undoubtedly has particular sequelae that are at present but obscurely understood.

These observations do not by any means exhaust the special influences to which extreme plus deviates are liable, simply by reason of intelligence. They do, however, serve to suggest that the eccentricities and lapses from common sense, popularly ascribed to genius, may be actual phenomena, and that they may arise from the psychological isolation to which extreme deviates are subject by reason of the extreme infrequency of like-minded persons. That very highly intelligent persons succeed as often as they do in achieving ' good ' personality under special difficulties suggests that great intelligence carries with it compensating functions that help to overcome hazards. It may be mentioned here that *self-sufficiency* and the tendency to *dominate* situations rather than to submit to them (36) may be such compensating factors. Furthermore, it is a matter of common observation that highly intelligent individuals have a keen sense of humor, typically, and are thus endowed with a ' saving sense.'

V. The Concept of Optimal Intelligence in Personality Development

From study of the eccentric influence of deviating intelligence, at both extremes (20, 35), the concept arises of *an optimal intelligence* that is neither minimum nor maximum. The problem for social psychology is to fix the limits of this optimum in terms of IQ. It is a situation in which there can be either too much or too little of a good thing (' good,' that is, for interaction with other persons, and hence for personality adjustment).

From observations recorded in the literature to this time (9, 35), the hypothesis may be offered that optimal intelligence, in the sense just referred to, lies above average, rather than at or below average, and centers approximately from 3 to 5 PE above the mean — that is, between 130 and 150 IQ (Stanford-Binet). Within this range, the person comprehends more clearly, but not too much more clearly, than the majority of his fellow men, and can thus get himself accepted as a supervisor and leader of human affairs generally, with accompanying emoluments and privileges (9). His vocabulary, his interests, and his hopes have, at this point, still enough *in common* with his contemporaries to enable and warrant coöperation. Beyond this range, however, mutual rejection begins to appear between the deviate and nearly all his contemporaries.

VI. Effects of Total Personality upon Intelligence

That intelligence can be affected in its caliber by other elements of personality seems unlikely. That it can be affected in its function is clear, both from observation and experiment. To cite a recent study, Simmons (53) found a large amount of overlapping between dull children, on the one hand, and highly intelligent children, on the other hand, in the ease with which they could be made to change their minds under suggestion applied from the outside. Although the high deviates were more self-sufficient, as a group, and less easy to dominate, still there were among them individuals whose intelligence *functioned* like that of dull persons, under suggestion, because of the total personality. The work of Collmann (15) already described in Chapter III of this volume shows how far intelligence is from being all powerful as a determiner of psychogalvanic reactions.

In the contributions immediately following, the discussion of these interrelations is amplified, in special connections. A large field for research is clearly projected from these discussions as a whole.

II

INTELLIGENCE AND PERSONALITY AS REVEALED IN QUESTIONNAIRES AND INVENTORIES

Irving Lorge
Research Associate, Institute of Educational Research
Teachers College, Columbia University
New York, New York

I. Types of Rating Scales and Inventories

The variety of the concepts of personality makes difficult the interpretation of its relation to intelligence. Allport (3) gives fifty definitions of personality, most of which include within their scope some aspect of dynamic organization, integration, or interrelatedness of a person's psychophysical adjustment in terms of attitudes, interests, values, and habits. Only recently have educators and psychologists attempted to go beyond haphazard observations, subjective inference, and intuition in their attempts to arrive at more precise and scientific understanding of personality.

The first type of approach was that of rating of personality traits — that is, the rating of behaviors, appreciations, and temperaments in terms of adjectival continua — by judges who were presumed not only to know how to rate but also to know the persons whom they rated.

Unfortunately, many of the values, attitudes, and behaviors to be rated were not evident in overt conduct. When the behavior designated as 'personality' could be observed through overt performances and actions, there was a reasonable basis for rating; when the behavior, however, did not manifest itself in overt action, a large subjective element entered into the rating of personality traits.

In order to get at the less overt manifestations of behavior, *self-rating* procedures were developed on the assumption that the subject may know aspects of his behavior that judges can only infer. Self-rating scales, however, had inherent in them many of the weaknesses of the rating scale.

A third type of approach was made through the self-analysis afforded by personality inventories, or schedules. In such schedules, the person checks whether he has or has not a certain behavior, or attitude, or opinion, or interest, or value in terms of 'Yes-No,' or of 'Agree-Undecided-Disagree,' or of checks on a graphic scale, or of choices of action. Such self-revealing inventories, if well-made, have advantages of objectivity, consistency, reliability, and reference norms, even if their validity is still to be established. It is validity that is most questioned in reference to self-evaluating inventories of the personality functions. The self-rating scale may not have in it the 'halo' of the judge, but it may have in it the halo of the cultural demand for the socially acceptable behaviors. To avoid the intentional or unconscious coloring of responses to personality inventories, various devices have been introduced, such as the word-association test, in which the subject has less likelihood of knowing the culturally acceptable pattern, and the performance test, in which the stress of the situation is such as to evoke the characteristic, or habitual, behavior of the individual.

All in all, more than a hundred tests, scales, schedules, or inventories have been developed or used to measure some function or functions of personality or its by-products. These many scales are oriented in many different understandings of personality. Some of them stem from the psychopathological referents of Freud, Adler, and Jung; others, from the empirical referents of Thurstone, Thorndike, and Pressey; but most stem from a combination of the psychopathological classificatory schema and empirical results. Each scale or inventory sets the concept of personality or personality function it seeks to test; that is, the scale, in a sense, operationally sets the meaning of personality that it seeks to rate. Thus, some attempt to get at manifestations of psychoneurotic disposition; some to measure tendency to aggression, or its opposite, submission; some to assay extroversion-introversion; others attempt to get at manifestations of behavior, such as cheating, or suggestibility, or leadership; still others, at moral knowledge and judgment.

Recognizing that schedules or inventories may be directed toward rather narrow aspects of personality, and recognizing that these aspects vary from scale to scale and from author to author, the investigation of the relationship of personality to intelligence is largely limited to the scale used. With this limitation, a fairly systematic review of the literature yielded some 200 correlation coefficients between intelligence and some scale of personality function.[1]

In order to utilize these correlation coefficients for the purposes that concern us, certain additional assumptions must be made. First, it is assumed, for purposes of generalization, that each measure of intelligence used is approximately equivalent to every other as a measure of intellectual ability. This assumption is not untenable when it is recognized that all reported studies of intercorrelation among measures of intelligence show a high degree of community, or, to put it another way, all intelligence tests are full of *g*.

Second, it is assumed that generalization can be made from the kind of populations studied. This assumption is less tenable because so many of the intercorrelations are reported for selected populations, like college students and, somewhat less often, high-school students.

Third, it is assumed that no bias operated in the selection of the correlation coefficients to be analyzed. This assumption is tenable because the search of the literature was systematic, although, unfortunately, many studies that report incidental correlations of personality to intelligence could not be searched for the evidence they might have yielded.

II. The Results

1. Correlations in General

With these limitations and assumptions, it is possible to report the evidence. The correlations of intelligence with measures of personality range from +.79 to —.49 with a median at +.04. Since many of the personality measures, however, report scores in terms of poorness of personality, the primary concern should be with the absolute size of the coefficient rather than with its sign. Half the correlations, on the basis of absolute size, range between .00 and .15 and only one-fourth

[1] I am indebted to Mrs. Helen H. Davidson of Teachers College, Columbia University, for her assistance in searching the literature for correlations between personality schedules and intelligence.

of them are greater than .30. It would seem, then, that some correlation between intelligence and personality actually exists. Some of the correlation will be introduced by the factor 'reading,' which is common to the two measures, and part may be the expression of a true relation between the measures.

2. Correlation between Intelligence and Different Measures of Personality

It is revealing, however, to consider the correlations by classifications of the measure of personality. The prototype of measures of *psychoneurotic tendency* is the Woodworth Personal Data Sheet. This data sheet was originally based upon descriptions of behavior of neurotic patients. Questions were selected to which the subject was asked to respond 'Yes' or 'No.' For instance, in the Woodworth Data Sheet are such questions as: "Are you afraid of noises at night?" "Have you ever fainted away?" Many variants of the data sheet have been developed, such as the Mathews and Cady revisions, the Laird Personal Inventory B2, Thurstone's Personality Schedule, Bernreuter Personality Scale B1-N, and others. More than 40 intercorrelations of such scale scores with intelligence have been tabulated, showing a range from +.18 to —.43 with a median of .00. Regardless of sign, half the correlations lie in absolute value between .00 and .06 and a fourth of them are greater than .12.

The prototype of the *introversion-extroversion* scale is based on Freud's collection of items from the writings of Jung. Heidbreder made a self-rating scale of the items by requiring the subject to rate himself for each item as 'plus,' 'doubtful,' or 'minus.' The student thus responds to such questions as whether he blushes, is self-conscious, daydreams, prefers reading to experience, and the like. Variants of the scale have been developed, as the Laird Personal Inventory C2, Bernreuter Personality Inventory Scale B3-I, Guilford and Guilford Nebraska Personality Scale, and others. For ratings on this type of scale the correlations of the scores with intelligence range from +.32 to —.19, with a median at —.01. Regardless of sign, half the correlations lie between .00 and .10 and a fourth above .13, absolute size.

The *ascendance-submission* scale is based upon self-ratings of conduct showing domination or submission. The best-known scale of this aspect of personality is the Allport A-S Reaction Study. Variants have been developed by Bernreuter, and by others in consolidated scales.

The few tabulated correlations with intelligence range from +.16 to —.15.

If all the correlations between intelligence and psychoneurotic tendency, such as introversion-extroversion and ascendance-submission, are combined with the correlations between intelligence and personality manifestations, such as feelings of inferiority, self-sufficiency, elation, and character, the range of the correlations between intelligence and these various aspects of personality is from .00 to .43, in absolute value, with half the correlations falling between .00 and .10, a fourth between .10 and .16, and the remaining fourth between .16 and .43. It can be seen that the schedules purporting to measure emotional tendency or behavior show relatively small correlations with intelligence.

The tests and inventories that have not yet been considered here specifically are the word-association type, the performance type, and the knowledge type. The tests of the *word-association* type have been developed as indirect measures of personality tendency. In the test the subject is given words to which he reacts by association. The Kent-Rosanoff has been scored to give number of pathological responses. The Pressey X-O test, a variant form of the word-association test, has been scored for 'affectivity' or 'richness in emotional association.' The correlation reported for Kent-Rosanoff with intelligence is +.31; those reported for Pressey X-O range from +.07 to —.18. Chambers, however, who rekeyed the Pressey to discriminate between good and poor students, reports a correlation of the rekeyed Pressey X-O with intelligence of +.53. In so far as richness of word association is positively regarded, word-association tests may be expected to reflect a weighting for knowledge, and thus, albeit indirectly, a correlation with intelligence.

It should be pointed out, parenthetically, that related to the word-association methods are the tests for investigating the total personality by means of unlearned or unstructured stimuli. The Whipple Ink Blot Test, the Rorschach figures, and the Morgan-Murray Thematic Apperception tests, each allow for an estimation of intelligence. For instance, on the Rorschach, the number and quality of the whole answers, of movement answers, of original answers and variety of content are considered symptomatic of intelligence. It is hoped that experimental evidence on this point will soon be reported in the literature.

Tests of *moral judgment or knowledge* are, to a marked degree, tests of information. As a matter of fact, one of the subtests of the

Army Alpha is a measure of ethical discrimination. It is well known that information and intelligence are positively correlated. Tests of ethical discrimination, of social intelligence, of moral judgment should, and do, give significant and positive correlations with intelligence, ranging from +.15 to +.65, or better.

Tests of *personality performance* have been developed with reference to getting objective, paper-and-pencil measures of tendency to cheating, resistance to suggestibility, manual movement, and the like. A host of experimenters, like Cady, Otis, Raubenheimer, Downey, Hartshorne, May, and Maller, have made tests of performance for some personality traits. The range of the correlations is from —.49 to +.77, with half the correlations lying in absolute value between .00 and .30, a fourth between .30 and .47, and the remainder above +.47. Tests of personality related to knowledge show substantial correlations with intelligence. It is not surprising, therefore, to find substantial correlations with intelligence in some attitude and interest schedules.

The distinction in the size of the correlations with intelligence and personality scores derived from the psychoneurotic, or introversion-extroversion scale as contrasted with the correlations of intelligence and measures of personality from word-association, performance, or knowledge tests may be due in part to the fact that the former are measures of single or narrower traits, whereas the latter get at some aspects of the total personality. Again, it may be due to the fact that different psychoneurotic, dominant, or inferiority behaviors or manifestations may be functions of different levels of intelligence. If personality is conceived as the adjustment that individuals make to their biological, psychological, and social environment, then intelligence will play a rôle in such adjustment. From one point of view, personality may be considered a psychophysical adjustment to the culture. In so far as items of behavioral adjustment are selected heterogeneously from all items, the correlation of total scores from such scales with intelligence must approach zero as a lower boundary. It seems reasonable to suppose that response to items may indicate specific maladjustments of behavior that only a bright child could make, or, conversely, that some responses are indicative of the environmental insensitivity of the duller child. Certainly, fantasy, imagination, and creative behavior are expected of the intellectually superior children. Yet, admission of that fact by the response ' Yes ' to the question: " Do you live in a make-believe life in addition to your real life? " would be scored as a tendency to maladjustment. In reference to the duller child, questions are asked to which he cannot respond. It would be surprising if a dull child had ever been concerned with such a question as: "Are you bothered by a feeling that things are not real?" It seems that, before any definite answer can be given to the relation of personality to intelligence, there is clearly a need to demonstrate the relation of each item of the inventory to measured intelligence. From another point of view, the relation of intelligence to personality as a function of age, developed by Furfey, suggests that the relationship of certain behaviors may be a function of the combination of chronological age and mental age.

The correlations between intelligence and measures of personality may underestimate the rôle of intellect in personality. It is suggested that personality is more complex than the scores on inventories are likely to indicate. While it may be statistically desirable to obtain scores on traits independent of intelligence, psychologically it is more important to get meaningful scores that will enable the educator to recognize and consider the interaction of intelligence with behaviors, values, and interests in our culture. Whenever the paper-and-pencil test measures some performance or some judgments of value related to personality, some significant correlation with intelligence may be expected. If such tests are scored for goodness or for social acceptability, the correlations will be positive. If psychologists are responsible for one generalization, it is that all positive traits are correlated positively. The evidence for this generalization is overwhelming. Intelligence is a positive trait, and good adjustment is also a positive trait. It seems reasonable to infer that, whenever personality is measured in terms of goodness of adjustment or of integration, a positive correlation will be the result.

Intelligence will not account for most of the variance in personality or personality adjustment. It is believed, however, that to understand the personality of individuals, their behaviors should be considered in relation to the particular subpopulation of which they are members by virtue of age, intelligence, and culture, because the subpopulation to which a person belongs tends in some degree to determine his habitual modes of reacting in his culture and environment. Intelligence, then, must be considered as one of the determiners, but not the only one, of personality.

III

INTELLIGENCE AND EMOTIONALITY

Frederick H. Lund
Professor of Psychology
Temple University
Philadelphia, Pennsylvania

I. Interdependence of Intelligence and Emotionality

That the processes represented in the concepts of intelligence and emotionality cannot operate in independent fashion is scarcely open to question. The complexity of these processes and the essential unity of body functions precludes any such possibility. Moreover, when viewed from the standpoint of current description and identification, it is evident that the connection between these two spheres of activity must be a very intimate one.

In the first place, the distinction between the central and autonomic nervous systems, considered basic, respectively, to the intellectual and emotional processes, is largely one of convenience and does not actually imply either anatomical or functional independence. In the second place, there is very little of our mental life that does not have its emotional overtones, its emotional effects, just as these in turn rarely occur in the absence of appropriate cortical representation.

The interdependence of intelligence and emotionality is evident also in the tendency to view both in terms of behavior characteristics, rather than in terms of definite states, powers, or capacities. Thus, the intelligence of the individual is made to refer to his general functional level, to his general capacity, and to the ease with which he adapts himself to new and changing conditions of his environment. Emotionality, on the other hand, is made to refer to the more obviously dynamogenic aspects of behavior. So long as behavior is of the habitual, routine variety, it is not said to be emotional. It becomes emotional only as the overt and somatic expressions provide evidence of direction and reinforcement through changes on the visceral level, such as are typically induced by sudden changes in the general situation.

II. Evidence from Psychopathology

That intelligence is not an isolable function, but a quality or characteristic of the principal reaction systems, energized and directed by a variety of organic and emotional states, is nowhere more in evidence than in certain cases of psychopathology. Consider, for instance, the wide variations in mental and physical output associated with the manic-depressive psychoses, with melancholia, and with schizophrenia. That variations in alertness and mental capacity under these conditions depend upon fundamental changes within the autonomic system is seen from studies in metabolic rate and bodily tonus characterizing these forms of disorder.

In a study of several hundred cases, Henry (31) found "a definite relationship between basal metabolic processes and certain emotional states." High psychomotor activity and high metabolic rate were uniformly present in the manic, the agitated, and the elated states. Low output and low metabolism, on the other hand, characterized the depressed and apathetic states.

Similar observations have been noted in disorders of the thyroid (50) and adrenal glands, the secretions of which are important factors in individual and sex differences in emotionality. The sluggish, underdeveloped conditions of the cretin and the overactive condition of the hyperthyroid have long been known. In a study dealing with the relation between intelligence and metabolic rate in grade-school pupils, Hinton (33) reports correlations of .74 and .66.

III. Conditions of Facilitation and Inhibition

Certain writers (37, 43) have emphasized the disruptive and inhibitory features of emotional states; other writers (57, 65) have called attention to just the opposite features. Much would seem to depend upon the type and intensity of the emotional state. Intense fear and agitation may, indeed, have a paralyzing effect. On the other hand, no blur, confusion, or inadequacy need attend emotions of the pleasant and inspirational type. The writer and artist may be filled with 'the fine frenzy of doing.' "Emotion interferes with thought," says Washburn (65), only when the movements made in emotion are incompatible with the movements and attitudes essential to thinking. This will be most likely to happen when the energy set free by the glandular processes in emotions discharges into the diffuse and random movements of motor explosion. Emotion will aid thought when conditions favor the discharge of this energy into the maintenance of a steady innervation of the trunk muscles that is the basis of introspectively reported feelings of will, determination, activity, or effort, and that secured the steady influence of the idea or goal.

That performance on educational and intelligence tests varies with the attitudes and emotions of the testee has been noticed by Corey (16), Mackaye (46), and others. Important, also, in learning and mastery situations is the emotional value of the material, as shown by Carter, Jones, and Shock (13).

IV. Control

There is, as we have observed, a constant interplay of mental and emotional factors in the life of the individual. So far we have approached the problem from the standpoint of the emotions and their influence upon intellectual processes. We have noticed that they may not only give direction to mental content but may also affect the quality and level of performance.

But the influences operating in the reverse direction are no less important. The control of our emotional life and the nature and extent of its external expression is largely a matter of intelligence.

The significance of intelligence in determining overt expression may be seen in the marked behavior differences characterizing cases of mental impairment and mental deficiency.

That intellectual and, therefore, cortical development should be a factor in emotional control is evident from the relation the cortical centers bear to the subcortical. This relation is one of dominance or of ascendance. The functions of the subcortical centers, which include those responsible for the organization and distribution of emotional impulses, are normally held in check by the higher centers. When, however, this check is removed, as happens under the influence of alcohol, drugs, and anesthetics, emotional reactions become much less inhibited and restrained. Loss of control, as in hysteria and 'nervous breakdowns,' very often follows upon conditions (those of overwork, excess worry, and insufficient rest) that impair cortical functions. Evidently the higher centers are the first to fail under these conditions, as they are the first to fail under the influence of drugs and anesthetics.

Impulsive and unrestrained behavior may be expected in the child, not merely because he is untrained and immature, but also because subcortical development is much more advanced than cortical. Similar unrestrained expression may also occur in adults who have failed to mature mentally.

A number of writers have given attention to the association of intellectual insufficiency with delinquency and to the likelihood that lack of control, owing to mental inadequacy, is responsible for these and other forms of social failure. De Greeff (27) notes that, while intellectual insufficiency does not in itself mean potential criminality, yet, from a study of many records, such insufficiency was found to be "one of the necessary prerequisites in 80 percent of moral crimes, acts of violence, and theft." In a similar study of delinquency, Armstrong (5) found that the average IQ of various offending groups is well below that of the general population. It is generally admitted, however, that in all such cases other correlates of low intelligence, such as poverty, poor environment, and insufficiency of training or opportunity, are more important contributing factors than is intelligence as such or lack of control.

In any account of the rôle of intelligence in relation to emotionality it should be kept in mind that these negative and restraining features, to which attention is typically called, are not of sole or even of primary importance. On the positive side, and particularly in the development of new avenues of expression, intelligence must be accepted as the guiding and molding influence.

With increasing maturity there should be not so much a loss of emotionality as a change of expression and an increase in the objects and situations that have emotional value. The mentally impoverished individual is apt to become emotionally impoverished, lacking, as he does, the resources that would enable him to cultivate new outlets for his emotional life, and particularly those provided in recreational activities and in the vicarious media of literature, art, and the stage.

IV

RELATIONS BETWEEN INTELLIGENCE AND MORBID ADDICTIONS

S. D. Shirley Spragg
Instructor in Psychology
Barnard College, Columbia University
New York, New York

I. Nature of Morbid Addictions

Addictions to ethyl alcohol or to one of the several derivatives of opium are by far the most frequent kinds of addiction found in our culture. In both, continued use of the drug induces a craving for the substance and a physiological need for its continued administration. The many social and therapeutic problems presented by the addict are too well known to need elaboration here. Recognition of their widespread importance has inspired a long list of studies on such aspects of the problem as the pharmacological nature of addicting drugs, the mechanisms of addiction, and a few on the characteristics of persons who become addicts.

Investigations of persons who become addicts have usually sought to discover whether certain intellectual or personality characteristics predispose them to addiction. Or, to state it as a question: Are persons who are introverts, or schizoids, or feebleminded, or 'only' children, or what not, more apt than are other persons to become addicted to habit-forming substances?

It is the purpose of this brief review to assess the available evidence concerning the possible relationships between intelligence level and the incidence of morbid addictions.

II. The Intellectual Characteristics of Addicts

The problem of the intellectual characteristics of persons who are chronic users of addicting drugs has been, and is, attended by considerable confusion and contradictory points of view. The moralistic implication is frequently encountered that persons who become addicted thereby demonstrate a deficiency in intellectual equipment; their addiction proves them to be *ipso facto* mentally inferior persons. Statistics showing the incidence of feeblemindedness among institutionalized alcoholics and drug addicts have at times been offered in support of this thesis, but since these data have rarely, if ever, included information about such variables as factors of selection and control, the results have contributed but little to an understanding of the problem.

The literature relevant to the problem of correlations between intelligence and morbid addictions is not extensive. There are, to be sure, a number of studies demonstrating that subjects, given experimental doses of alcohol or other drugs, do more poorly on various mental tests and on intelligence tests than do control subjects or the same subjects during control periods. The early studies of Dodge and Benedict (20) demonstrated deleterious effects of small doses of alcohol on a variety of intellectual functions. Hollingworth's experiments (34), as well as those of other workers, have provided corroborating evidence in several experimental settings, and Cattell (14) found that 20 grams of alcohol definitely lowered the intelligence quotients of his subjects.

although 10 grams had practically no effect. The experimental work in this general field has been ably reviewed by Darrow (18) and by Miles (47). Comparable evidence on the effects of experimental doses of opiates is not available, but clinical findings indicate that psychometric tests given at such times would very probably show a decrement, perhaps of the same order as that following doses of alcohol. But such evidence, although valuable for many purposes, throws little light on the problem of relations between intelligence and addiction.

Among the few early studies directly related to the present problem is that reported in 1925 by Kolb (40). This investigator first reviewed the previous studies and found only two deserving of citation. In one of these, Anderson [1] had examined 70 cases of narcotic addiction brought to the Boston Municipal Court and found (using an old form of the Binet test) that 54.5 percent of them had an M.A. below twelve and that 28.5 percent were feebleminded. The other source cited was Jewett,[2] who examined 200 narcotic 'repeaters' at Bellevue Hospital and found that 25 percent had IQ's below 70, 40 percent between 70 and 80, 30 percent between 80 and 90, with only 5 percent having 'normal' intelligence. Jewett opined, however, that if every addict were studied, the 5 percent would be raised to a much higher value. Kolb, evaluating these two studies, states (40, p. 165):

> It is obvious that the surveys cited above were made on two selected groups having intelligence lower than that of the average addict, and for this reason the results obtained cannot be accepted as a true index of the relation of intelligence to the etiology of addiction.

Kolb gave intelligence tests (Stanford-Binet) to 100 narcotic addicts as part of a general psychiatric survey of addicted persons. In general, the cases tested were selected just as they came to him, but not every addict in the group was tested; thus, by the time 100 addicts had been tested, 150 had been seen. That a selective factor was clearly present here is indicated by Kolb's statement:

> The results do not, therefore, represent a true cross-section of the intelligence of addicts in my series of cases. When lack of time was a factor in the examination of several cases so that not all could be given intelligence

[1] V. V. Anderson, *Boston Med. & Surg. Jour.*, 176, 1917, 755–757. Cited by Kolb (40, p. 163).

[2] S. P. Jewett, *Tice's Practice of Medicine.* (W. T. Prior: Hagerstown, Md. Vol. 8) p. 21. Cited by Kolb (40, p. 163).

tests, those who obviously had a high grade of intelligence were not tested . . . of 20 professional men who were examined at their home, only two were given intelligence tests. (p. 164)

It is therefore not surprising that Kolb's subjects, as a group, tested somewhat lower than a sample drawn from the general population would be expected to. His results show a mode between 86 and 95 IQ with ten scores below 70 and only one above 110. From these findings he concludes as follows:

. . . judging from these 100 cases, feeblemindedness or mental dullness is apparently not of itself an important predisposing cause of drug addiction. In fact, some of the worst of them — from the standpoint of trifling reasons given for their original addiction and for relapse after cure — had intelligence quotients around 100. (p. 167)

In a further study, Kolb (41) classified morphine addicts into five groups on the basis of causes of their addictions. It is significant that feeblemindedness or low-grade intelligence is not listed as a predisposing factor. Kolb found it neither necessary nor valid to provide a class for those becoming addicts because of deviations in intelligence level.

Wall's investigations of male and female alcoholics (63, 64) included many cases of superior education and training — college graduates, musicians, artists, and executives.

Hall (29) studied the mental and physical efficiency of 57 women drug addicts at the State Reformatory for Women in Illinois. She states that psychiatric examination showed them, as a group, to be neurotic or twisted personalities. Psychological tests (Army Alpha) were given after the acute symptoms of withdrawal had been passed.

Army Alpha I.Q.'s ranged from 67 to 127, or from the mental defective to the very superior level. The mean I.Q. was 96.2, with a standard error of 2.7. The standard deviation of the distribution was 16.5. The mean I.Q. of the addict group was close to the mean of the general population and very much above the mean of 83.9 of the Reformatory population. (p. 339)

East, an English Commissioner of Prisons, reports (23) that only a small number (7.4 percent) of the 698 mentally defective persons in his jurisdiction convicted of various offenses during the period 1921–1930 were chronic alcoholics. He concludes, after citing further evidence, ". . . the mentally defective person is seldom an alcoholist; if he becomes an addict, it is incidental to the mental defect." (pp. 158–159)

However, in a recent study, so far reported only in abstract, Dimmock (19) finds evidence of mental inferiority among drug addicts. This investigator gave the Babcock Mental Deterioration Test to 156 incarcerated male addicts, aged 18 to 68, with an addiction range of 2 months to 25 years. All subjects had been withdrawn from their drug for 6 months at the time of their testing. Comparisons of the mental efficiency of this group were made with two control groups of like age and ability. The results were said to indicate statistically reliable differences adverse to the experimental group in all comparisons with the normal groups. Further, Dimmock asserts that there were no reliable differences in mental efficiency related to the duration of addiction, when age and mental ability were held constant. He suggests "that the functional characteristics of the drug addict revealed by these tests existed prior to drug addiction and that they may have contributed directly or indirectly to the formation of the drug habit."

In a recent report, Pescor (49) has presented material gathered from a study of 1,036 patients admitted for treatment of narcotic drug addiction to the United States Public Health Service Hospital at Lexington, Kentucky, during the year 1936–1937. The majority of the patients were given the Army Alpha and Stanford-Binet psychometric tests. Certain others were given the Army Beta, the Pintner-Paterson, or other performance tests:

> However, for the sake of uniformity, scores were all converted to mental ages, the average mental age being computed as 13 years and 8 months. On the basis of 15 years [*sic*] as the average adult level of intelligence, the present subjects are subnormal . . . approximately 8.7 percent of the present subjects have an I.Q. below 70, and 83 percent an I.Q. above 75. (p. 17)

In connection with the findings of the last two investigations, it seems imperative here to raise a question that is relevant to all results pertaining to addicts reported from institutionalized (incarcerated) populations; that is, what degree of selection does the fact of being in an institution represent? May it not very well be that the most intelligent addicts are those most apt to be successful in avoiding arrest and conviction and in maintaining their sources of supply without coming to the attention of authorities, so that they do not appear at all in surveys of addicts? If this assumption be valid, and available evidence seems to confirm it, then all intelligence-test scores reported for groups of institutionalized addicts are spuriously low by an undetermined, and possibly undeterminable, amount.

Such results also contribute little if anything to our knowledge of the etiology of drug addiction. To show that institutionalized addicts exhibit intellectual deficiencies when tested by various psychometric procedures does not, of course, establish anything with regard to causal factors operating. Does low-grade intelligence *cause* alcoholism or morphinism? Or, are the poor scores made by institutionalized addicts a *result* of physical, neural, and personality changes produced by the chronic intoxication of the drug itself? Cause and effect relationships here can never be isolated by the mere amassing of addicts' scores on psychometric tests. If we possessed test scores of addicts, made *before* they become addicted, that would be another matter. Such data would be highly valuable, but nothing short of a universal testing program would provide the information desired.

III. Conclusions to Be Drawn from the Literature

1. The material considered in this review of the literature forces the general conclusion that deviations in intelligence (as measured by the standard intelligence tests) have not been shown to be related to proneness to addiction. Low intelligence has not been shown to predispose one to addiction, nor high intelligence to protect one from it.

2. The studies reviewed convince the writer that there is a paucity of accurate knowledge about the mental characteristics of drug addicts, since there has so far been no study that has measured a truly representative sample and compared the results for this sample with results for control groups drawn from the same types of the general population from which the addict cases come.

3. Causes of addiction may more profitably be sought in non-intellectual characteristics of the individual than in the factors measured by conventional tests of intelligence. Personality deviations of many kinds, with emotional conflicts, feelings of inferiority, attempts at flight from reality, and so forth, all seem to be much more relevant factors in the etiology of morbid addictions than are the scores that addicts make on intelligence tests, or even the scores that they presumably would have made if they had been tested prior to their addiction. Addictions are symptoms rather than disease entities; and they are symptoms, not of intellectual inferiority, but of personality maladjustments.

V

INTELLIGENCE AND DELINQUENCY

Harold M. Williams
Supervisor for the Mentally Handicapped
State Department of Public Instruction
Madison, Wisconsin

The analysis of the relationship between intelligence and delinquency involves three major problems: (1) the validity of the data available, (2) the distribution of intelligence among delinquents, and (3) the evaluation of hypotheses regarding the nature of the relation.

I. Sources of Error Affecting the Validity of the Data

1. Degree of Rapport Secured in Testing

Delinquents are ordinarily tested during the period of court appearance or after commitment. Under such circumstances the subject's attitude and general efficiency are probably seldom at their best. Canady (12) found a small systematic difference just by having Negro children tested by a Negro examiner, for example. Inadequate rapport would systematically lower the scores of delinquents.

2. Choice of Basic Age

The basic age used has not been consistent. Slawson (55) has pointed out that using 16 years as the basic age caused 31.1 percent of his group to fall into the feebleminded class, whereas a basic age of 14 reduced this percentage to 15.4. Lane and Witty (42) specifically mention this point also. This criticism applies particularly to the early studies. This variation in usage would of course operate to reduce the agreement between studies.

3. Lack of Control Groups

Many studies, especially the earlier ones, have been made without control groups other than the one furnished by the standardization of the test itself. Some of the later workers, however, have exercised considerable care in this respect. In the present problem it is necessary to establish clearly the need, or the lack of need, for special control groups; the data on this point are presented in the third section of my discussion. On the basis of the evidence given there, it is concluded that the lack of comparable non-delinquent control groups has systematically made the delinquent distributions appear in an unjustly unfavorable light.

4. The Suitability of Verbal Tests

Several writers have raised the issue of the adequacy of the more verbal type of test with the delinquent. Slawson (55) concluded that the inferiority of the delinquent was most marked in verbal abstract intelligence. He found that delinquents tended to be on a par with non-delinquents in mechanical aptitude. Aden (2) found a low correlation between the Pintner-Paterson and the Stanford in a group of juvenile court cases in Los Angeles County. Selling and Stein (52) also raise this issue in a vocabulary study in which they compared the achievement of a group of St. Charles School delinquents with that of a high-school group from an area of low delinquency. They conclude that the delinquent tends to come from an inadequate cultural background and that this may handicap him on a test of intelligence involving a large verbal element. Knight (38) found that the discrepancy between scores on the Arthur and the Stanford tests was greater for delinquents than non-delinquents. The Arthur scores were the more favorable in the case of the delinquents.

This factor would operate systematically against the intelligence rating of delinquents if the more verbal type of test were used as a criterion.

5. Problems of Sampling

One of the most intricate problems appears in connection with the selection of a delinquent sample. The most commonly used criteria of delinquency have been court appearances and institutional commitment.

The first criterion is subject to the caprice of anyone who wishes to file an information. In a recent bulletin of the Children's Bureau (62), 37 percent of such complaints were on the ground of mischief and of being ungovernable. A wide range of behavior is possible under such categories. Stealing, running away, and truancy, which account for another 55 percent, would seem to be more objective.

Institutional commitment should be a more valid criterion of delinquency than court appearance, since such cases have at least been adjudged according to the rules of evidence. Institutional commitment, however, being the alternative to probation, varies according to many adventitious circumstances, such as the availability of suitable probation officers. What may occur to the distribution of intelligence in an institution on the basis of a changed policy of commitment is shown by Sullivan (59). Whereas in 1918 the percentage of feebleminded in the Whittier School was 29.9, by 1926 it had been reduced, as the result of a declared policy, to 2.2 percent. Analogous changes were evident at all other intelligence levels also. Similar changes have probably occurred elsewhere without being noted in the studies now available for this analysis.

Two general statements seem to be justified by the foregoing analysis: (1) delinquent intelligence has probably appeared in a systematically unfavorable light, and (2) the validity of the results may be considerably impaired by sampling impurities.

II. The Distribution of Intelligence among Delinquents

An excellent summary of the results of studies prior to 1923 has been given by Pintner (50). Lane and Witty (42) have presented a tabulation of results up to 1935. Results from some additional studies will be presented here.

1. A Systematic Shift in Results

Pintner noted one very interesting fact regarding the reports then available from the preceding twelve years. While the median for 16 earlier reports was 64 percent of feebleminded, the median for the 16 later reports was 26 percent. He held that much of the variation was due to differences in the mental make-up of the group examined. Miner (48), however, in 1919 questioned the validity of the results of these early studies on the ground that they made too broad an interpretation on too narrow a criterion. Lane and Witty (42) interpret these early findings as being due to the use of incompletely standardized adaptations of the Binet test, and to variations in the basic age used. The writer believes the change during this period to be due in part at least to changes in policy of dealing with cases. This change may have been effected by two factors: first, a growing clarification of the concept of intelligence at this time; and second, changes in policy of dealing with delinquents consequent to this classification and to Flexner's report (24) on juvenile courts.

2. Proportion of Feeblemindedness among Delinquents [1]

The general tendency of recent studies has been to make Pintner's later median the maximum. Burt (10) reported 8 percent feeble-

[1] In the studies here reported the criterion of feeblemindedness is a Binet IQ of 70.

minded in his sampling of 200 English delinquents. Slawson (55) found 31.1 percent among his 553 institutional boys. McCaulley (45) reports 16 percent among 100 boys. On the basis of Breguet's report (7) and statistics gathered from other sources, the percentage of defective delinquents in New York State institutions is estimated at about 40 percent. In the Iowa Training School, 11.8 percent of 279 boys were feebleminded.[1] Growdon and Calhoon (28) reported 6.8 percent among 1,104 admissions to the Boys' Industrial School of Ohio and 17.5 percent among 313 admissions to the Girls' Industrial School.

3. Central Tendency

In the estimates of central tendency a distinct tendency toward agreement is found in the later studies. Lane and Witty (42) report a Binet average of 79; Slawson (55) reports 86.8, Burt (10) 89. Aden (2) reports a mean IQ of 85 on 410 boys and girls, court cases in Los Angeles. An approximate mean calculated from Breguet's (7) data gives 90.0. The mean value for the Iowa group was 79.0.

The proportion of superior intelligence among delinquents has uniformly been reported as very small. Slawson (55) found 11.8 percent; Sullivan (59), 12.2 percent; Burt (10), 4 percent; Lane and Witty (42), 4 percent; Breguet (7), 9.5 percent. In the Iowa study the percent was 3.0.

The more recent studies have yielded, therefore: (1) from 10 to 30 percent feebleminded, (2) a central tendency of about IQ 85, and (3) a very markedly reduced proportion of superior intelligence (3 to 12 percent). Though these results are much higher than the early estimates, they range far below the general population, especially at the upper levels. There seems to be little difference between court and institutional cases.

III. Hypotheses regarding the Nature of the Relationship between Intelligence and Delinquency

1. Low Intelligence as a Direct Causative Factor

Goring (26) subscribed definitely to the view that 'natural stupidity' was a direct causative factor in criminality. This view was, and is, held by others. The principle usually applied is the presumably

[1] These data were obtained through the courtesy of Dr. H. M. Skeels and the Iowa Board of Control.

lower ethical discrimination of the dull. Terman (60) expressed the relation in terms of the ability to foresee and weigh possible consequences.

Several attempts have been made to demonstrate this view experimentally. Hill (32), however, found negative results from the comparison of 517 reformatory boys with 1000 high-school students and 148 adults on rating items of conduct. Weber (66) found little difference between 138 female delinquents and university women on the Brogan test of moral perception. Other studies have been more or less in agreement in yielding negative results.

2. Attempts to Scale Delinquency

Further insight may be gained, perhaps, from the studies that have attempted to differentiate degrees of delinquency. The principal criteria used have been recidivism and seriousness of offense. Durea (22) concluded that it was impossible to estimate the degree of delinquency from mental age, life age, or developmental age. Slawson (55) and Lane and Witty (42) found no relation between intelligence and number of convictions in court or seriousness of delinquent activity. Thus, the results of efforts to relate intelligence to degree of delinquency have been largely negative.

3. Low Intelligence and Delinquency as Correlates of Inferior Status

One may summarize the evidence up to this point as follows. *As sampled at present,* the proportion of feebleminded among delinquents is greater than expectancy on the basis of the standardization of the tests; the proportion of superior, much less. The typical delinquent is a dull normal. Attempts to differentiate the delinquent from the normal on the basis of understanding as measured by the ethical discrimination tests have yielded largely negative results. No clear differences in intelligence have resulted from attempts at differentiation of degree of delinquency.

It would appear that further hypotheses regarding the association of the two factors are called for. Before attempting this, it will be necessary to present certain further data. It has been well demonstrated that delinquents sampled in the mass come from inferior social environments. A good recent summary of this evidence has been made by Sullenger (58). Caldwell (11), for example, has shown that the occupation of parents of two-thirds of 492 institutional delinquent boys in Wisconsin was of the unskilled classification.

A much neglected, though, in the writer's opinion, crucial, line of evidence bearing on the present problem is the fact that it has been repeatedly shown that the intelligence of children of socio-economically inferior groups is also, with or without delinquency, inferior. Terman (61) long ago reported the correlation between IQ and occupation of the father (Taussig classification) as .48; he reported the median IQ of children from the two lowest classifications as 82.5 (150 first-grade children). Stoke (56) found a correlation of .30 with status in a study of 501 children. Goodenough (25) has also emphasized the variations in intelligence with varying status of the family.

Daniel (17), sampling Negro children in Richmond, found the average IQ of 100 delinquents to be 74.0, that of 100 behavior-problem children, 86.4, that of 120 non-problem children, 93.6. Lane and Witty (42) found differentiations within the delinquent group on the basis of probable socio-economic differences in the community; from a delinquency area, the mean IQ was 87.5, from town and country areas, 91.1. In Iowa, where the state maintains both orphanage and delinquent institutions, differentiations in court commitment may be made according to the fact of delinquency. A free transfer policy between institutions is also practiced. The distribution of intelligence of the older orphanage children given by Skeels and Fillmore (54) is almost identical with analyses made by the writer for both boy and girl delinquents.[1]

Among siblings, where there obtains a still higher degree of control, Ackerly (1) found the intelligence of 30 delinquents about the same as that of 30 non-delinquents. In what is perhaps the most crucial study now available, Healy and Bronner (30) found practically no difference between delinquents and their siblings, except in individual cases where intelligence was a definite factor in the total situation. Lane and Witty (42) specifically conclude that the intelligence of delinquents is roughly of the same order as that of non-delinquents sampled from comparable groups.

The writer feels, then, that the data now available must be interpreted as follows: Stated in its most conservative form, it is concluded that both data and hypotheses are lacking for the demonstration of a direct causal relationship between low intelligence and delinquency, except, perhaps, in particular individual cases. A more reasonable interpretation of the results from mass data is that samplings from

[1] From the records of the Iowa Board of Control.

inferior groups show a higher incidence of both delinquency and low intelligence.

The generalized quality of defect in the underprivileged is familiar to all who have had contact with them. The effects of environmental inadequacies ramify into an extraordinary variety of channels. Stoddard, Wellman, Skeels, and their group have presented very suggestive evidence in support of the functional view of intelligence,[1] a view that has perhaps been too easily taken for granted in delinquency. It may well be that underprivilege, while adversely affecting intelligence through inadequate stimulation, is at the same time disproportionately productive of conditions basic to delinquency.

VI

INTELLIGENCE AND MOTIVATION

Kurt Lewin
Professor of Child Psychology
Iowa Child Welfare Research Station
State University of Iowa
Iowa City, Iowa

Behavior can be said to be dependent upon ' ability ' and ' motivation.' To state this idea in somewhat less vague terms, one can say: what actually is done by the individual at a given time depends partly on the ' possibilities ' inherent in the person and in the situation, partly on the forces (drives, goals) acting upon the person at the time. The concept of intelligence is definitely related to the ' ability ' and not to the motivation; in other words, to the potentialities inherent in the person. The purpose of intelligence testing is to measure the ability for intelligent action separated from motivational factors.

The possibility of testing the intellectual ability of a person at a given time substantially independent of motivation does not, of course, mean that the growth of intelligence or even the momentary actual intellectual ability of a person (which may be defined as the best that he is able to do intellectually in his present state) is necessarily independent from motivation. The testing procedure emphasizes strongly the

[1] See especially: Part I, Chapter XIV; Part II, Chapters XX and XXVI in this Yearbook.

relation between intelligent behavior and motivational situation by requesting something approaching an optimal state of the subject during testing; the subject should be comfortable emotionally and should readily accept the goals required in the different tasks of the test.

Two problems might be distinguished in the relation between motivation and intelligence: first, the motivational situation might change the intellectual ability of the individual in this situation but not permanently; second, the motivational situation might change the intellectual ability of a person permanently either by a sudden, strong influence or by the existence of certain motivational situations during a long period, which might gradually hamper or speed up the development of the ability of the person.

These are the theoretical possibilities as to how motivation might influence intelligence. Whether such relations actually exist is a different matter. Psychology has come to realize that no one factor determining psychological behavior or development is entirely independent of the other factors. One might, then, expect mutual influence between intelligence and motivation. The nature and amounts of such influences would have to be determined by research.

Permanent changes in intellectual ability brought about by changes in the motivational situation of the person are described in reports that give, for example, a detailed story of the way the school achievement of an individual is basically changed by a change in his attachment to his mother and father and to their respective ideals. To my knowledge, however, no measurement of IQ has been carried through in these cases. In the cases reported by Wellman, Skeels, and others, the change in intelligence might be due partly to motivational factors.

Our knowledge about the relation between the momentary intellectual level and the motivational situation is rather limited, too, although some data are available.

As for effects in the reverse direction, doubtless the level of intellectual ability influences motivation as well as is influenced by it. What goal is chosen, under what conditions a goal is changed, and what path will be chosen to attain it depend partly upon how the individual perceives the situation. The degree of difficulty in attaining his goals obviously affects his persistency.

As to the dependence of the momentarily manifested intellectual ability upon motivation, we know from the experiments with animals and children (39) that the best intellectual achievement is obtained under what might be called 'optimal motivation.' If the interest is too low or if the zeal is too great, the intellectual achievement is reduced.

Of course, one can say that in such cases of emotional tension or overmotivation the person's ability did not change but only the psychological situation existing for him. Using operational definitions, however, and keeping in mind that the psychological environment and the person are but one dynamical field (frequently called 'personality'), one might as well say that as a result of the emotional stress the ability of the person within the same objective environment has become less. The differences between these formulations are but differences in terminology.

The empirical problem remains whether overtension in regard to a given goal affects only actions directed toward that goal or spreads out in such a way as to affect other activities and abilities. It is obvious that this question has a direct bearing on the scope and permanency of the influence of motivation on abilities. That is why it might be justifiable to mention here the study (6), in which the writer participated, of the effect of frustration on regression. The ability studied was constructiveness in play, which is closely related to intelligence, though perhaps not 'intellectual ability' in the sense commonly subjected to testing.

Children with mental ages between two and six years were observed during 30 minutes of free play in a standard situation. The constructiveness of the play, when rated minute by minute, showed a correlation with chronological age of .81, and with mental age of .77. Later the children were brought a second time into the same set-up (same toys, same placement), but this time one wall of the room had been changed into a partition through which far more attractive toys were visible to the child. The child first was permitted to play for a short time with these more attractive toys, but then had to return to the less attractive ones. Again the play was observed for 30 minutes and its constructiveness rated. Figure I gives the original constructiveness and the amount of change during the 'frustration' situation. The average decrease in constructiveness is equivalent to a regression of about 18 months. For the older children (mental age, 4.1 to 6.1 years) the decrements amount to 22 months in mental age; for the younger (mental age, 2.7 to 4.9 years) to about 9 months in mental age. One can measure the actual psychological frustration created in the standard situation within the different individuals by symptoms independent of constructiveness. By classifying the cases into those with weak and those with strong frustration, one finds that weak frustration leads generally to a small increase in constructiveness, whereas strong frustration leads to a definite decrease (Fig. II). These results are well in line with theoretical expectations. One of the main differences between a younger and an older child is the greater differentiation of the latter. This greater differentiation is proved by the greater variety of possible behavior, the greater variety of needs and emotions, and by other symptoms.

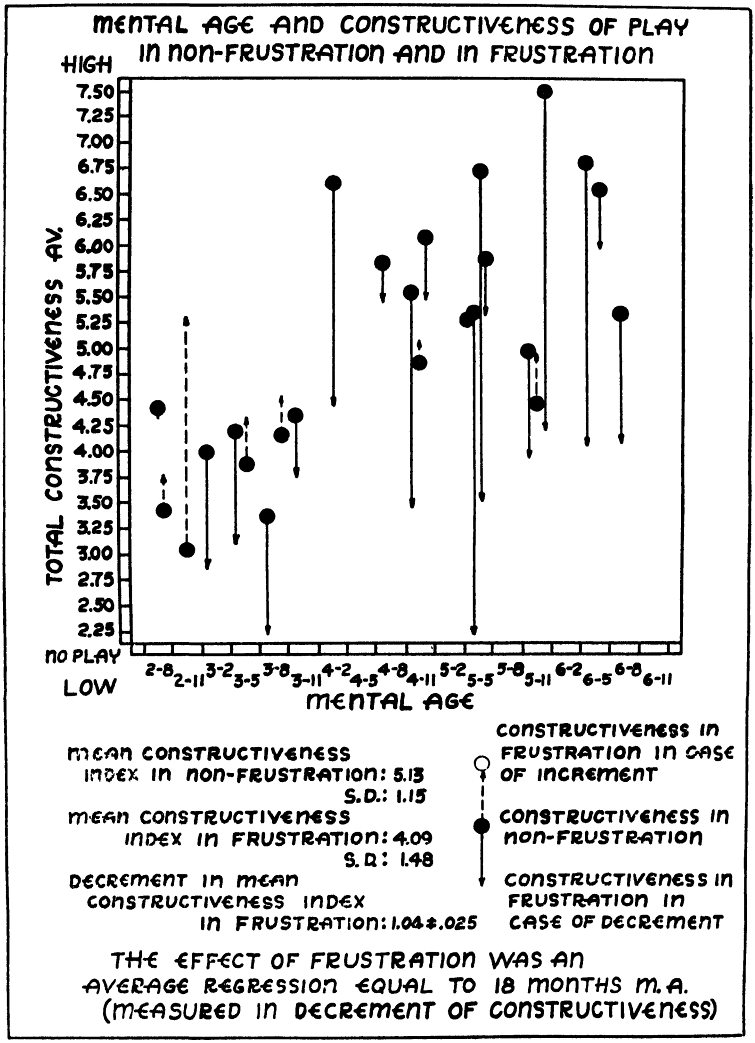

FIG. I. — MENTAL AGE AND CONSTRUCTIVENESS OF PLAY IN NON-FRUSTRATION AND FRUSTRATION.

Dynamically, a higher differentiation probably means the existence within the person of a greater number of systems that are (functionally) relatively separated. On theoretical grounds, it is to be expected that the degree of differentiation of a person at a given time is not independent of the pressure brought to bear upon him: if this pressure is above a certain level, only those personal systems will remain relatively independent that are separated by

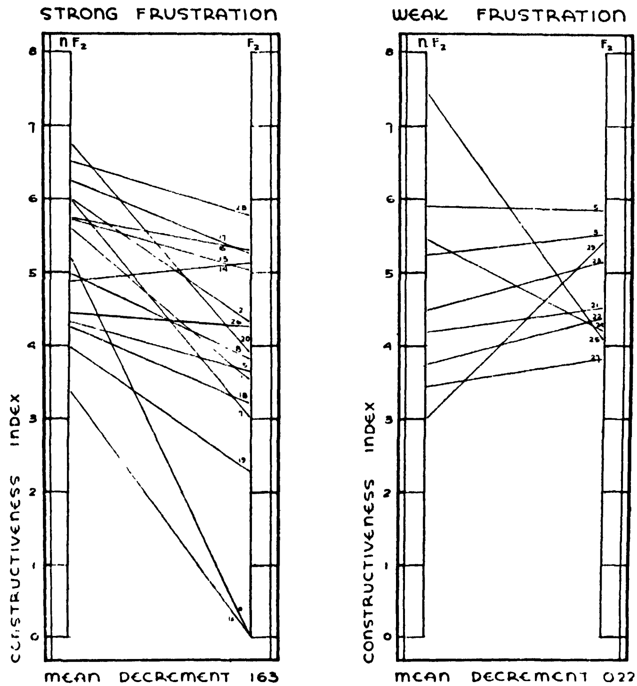

FIG. II. — EFFECT OF STRONG AND WEAK FRUSTRATION UPON CONSTRUCTIVENESS

rather strong dynamical walls. In other words, the differentiation of a child who is exposed to too great forces or is in a state of overtension becomes equivalent to that of a younger child; he will regress to a more primitive level.

On the other hand, a person will bring his total abilities into play and therefore reach his highest level of constructiveness only if he gets involved in the activity with his total personality, with the totalities of his systems. This seems to be the reason why weak frustration or a certain amount of difficulty leads frequently to better work: it is more likely that under this condition the total personality becomes involved. However, if the forces acting on the person go much above this level of intensity, the constructiveness will decrease as a result of the dedifferentiation or 'primitivation' (if one may invent a term) of the person.

It seems to me probable that what might be called the 'momentary intelligence level' of the person (that is, the maximal intellectual requirement he is able to fulfill at a given time) depends upon the strength of forces acting on the person in a similar way as constructiveness does. The correlations between constructiveness and mental age indicate this; in fact, intelligent behavior can be regarded as one type of constructive behavior. There seems to exist a close relation between intelligence and the degree of differentiation of the person. One would expect, therefore, the 'momentary' mental age of a person to depend upon the motivational situation, particularly upon the strength of the forces involved, the amount of pressure or tension.

Whether such influence will have a temporary or permanent effect will then depend upon the duration of this situation, the type of this influence, and at what state of development of the person the influence occurs.

References

(1) Ackerly, S. "Rebellion and its relation to delinquency and neurosis in 60 adolescents." *Amer. Jour. Ortho-psychiat.*, 3: 1933, 146–158.

(2) Aden, M. E. "Preliminary report of a study of the motor ability of delinquent boys and girls of the Los Angeles Juvenile Court." *Jour. Delinq.*, 10: 1926, 351–367.

(3) Allport, G. W. *Personality.* (Holt & Co.: New York, 1937) 588 pp.

(4) American Psychiatric Association. *Proceedings of the First Colloquium on Personality Investigation.* (New York City, December, 1–2, 1928) 102 pp.

(5) Armstrong, C. P. "Juvenile delinquency as related to immigration." *School and Soc.*, 38: 1933, 61–64.

(6) Barker, R., Dembo, T., and Lewin, K. "Experiments in frustration and regression in children." (Unpublished study)

(7) Breguet, R. "Preliminary survey of 1000 case histories of inmates of the Elmira reformatory." *Amer. Jour. Med. Jurisprudence,* 1: 1938, 101–118.

(8) Briggs, A. E. *The Concept of Personality.* (University of Southern California Press: Los Angeles, 1935. School of Research Studies, No. 7; Social Science Series, No. 10) 55 pp.

(9) Burks, B. S., Jensen, D. W., Terman, L. M., and Others. *Genetic Studies of Genius. III. The Promise of Youth.* (Stanford University Press: Stanford, Calif., 1930) 508 pp.

(10) Burt, C. L. *The Young Delinquent.* (D. Appleton-Century Co.: New York, 1925) 619 pp.

(11) Caldwell, M. G. "The economic status of families of delinquent boys in Wisconsin." *Amer. Jour. Sociol.,* 37: 1931, 231–239.

(12) Canady, H. G. "The effect of 'rapport' on the IQ; a new approach to the problem of racial psychology." *Jour. Negro Educ.,* 5: 1936, 209–219.

(13) Carter, H. D., Jones, H. E., and Shock, N. W. "An experimental study of affective factors in learning." *Jour. Educ. Psychol.,* 25: 1934, 203–215.

(14) CATTELL, R. B. "The effects of alcohol and caffeine on intelligent and associative performance." *Brit. Jour. Med. Psychol.*, 10: 1930, 20–33.

(15) COLLMANN, R. D. *The Psychogalvanic Reactions of Exceptional and Normal School Children.* (Teachers College, Columbia University, Contributions to Education, No. 469: New York, 1931) 88 pp.

(16) COREY, S. M. "The effect of motivation upon the relationship between achievement and intelligence." *School and Soc.*, 41: 1935, 256–257.

(17) DANIEL, R. P. *A Psychological Study of Delinquent and Non-Delinquent Negro Boys.* (Teachers College, Columbia University, Contributions to Education, No. 546: New York, 1932) 59 pp.

(18) DARROW, C. W. "Psychological effects of drugs." *Psychol. Bull.*, 26: 1929, 527–545.

(19) DIMMOCK, G. B. "The mental efficiency of the narcotic drug addict." *Psychol. Bull.*, 35: 1938, 644–645. (Abstract)

(20) DODGE, R., and BENEDICT, F. G. *Psychological Effects of Alcohol.* (Carnegie Institution of Washington: Washington, D. C., 1915) 281 pp.

(21) DOLL, E. A. "Social adjustment of the mentally subnormal." *Jour. Educ. Research,* 28: 1934, 36–43.

(22) DUREA, M. A. "Relationships of certain indicators of juvenile delinquency." *Educ. Research Bull.*, 14: 1935, 150–154.

(23) EAST, W. N. "Mental defectiveness and alcohol and drug addiction." *Brit. Jour. Inebriety,* 29: 1932, 149–168.

(24) FLEXNER, B., and BALDWIN, R. N. *Juvenile Courts and Probation.* (D. Appleton-Century Co.: New York, 1914) 308 pp.

(25) GOODENOUGH, F. L. "The relationship between socio-economic status and mental traits." [In] *White House Conference on Child Health and Protection, Report of the Committee on Growth and Development:* Part IV. *Appraisement of the Child. I. Mental Status. II. Physical Status.* (D. Appleton-Century Co.: New York, 1932) 344 pp. (pp. 178–188)

(26) GORING, C. B. *The English Convict.* (H. M. Stationery Office: London, 1919) 275 pp.

(27) GREEFF, DE. "Le niveau intellectuel et la criminalité." *Jour. belge de neur. et de psychiat.*, 33: 1933, 114–123.

(28) GROWDON, C. H., and CALHOON, C. H. "An analysis of cases admitted to the Boys' Industrial School and the Girls' Industrial School during the year 1931." *Quart. Bull. Assoc. Ohio Dept. Public Welfare,* 10: 1933, 6–14.

(29) HALL, M. E. "Mental and physical efficiency of women drug addicts." *Jour. Abnorm. and Soc. Psychol.*, 33: 1938, 332–345.

(30) HEALY, W., and BRONNER, A. F. *New Light on Delinquency and Its Treatment.* (Yale University Press: New Haven, Conn., 1936) 226 pp.

(31) HENRY, G. W. "Basal metabolism and emotional states." *Jour. Nerv. and Mental Disease,* 70: 1929, 598–605.

(32) HILL, G. E. "The ethical knowledge of delinquent and non-delinquent boys." *Jour. Soc. Psychol.*, 6: 1935, 107–114.

(33) HINTON, R. T., JR. "The rôle of the basal metabolic rate in the intelligence of ninety grade-school students." *Jour. Educ. Psychol.*, 27: 1936, 546–550.

(34) HOLLINGWORTH, H. L. "The influence of alcohol." *Jour. Abnorm. and Soc. Psychol.,* 18: 1923, 204–237; 311–333.

(35) HOLLINGWORTH, L. S. "Development of personality in highly intelligent children." *Yearbook Nat. Element. School Prin.,* 15: 1936, 272–281.

(36) HOLLINGWORTH, L. S., and RUST, M. M. "Application of the Bernreuter inventory of personality to highly intelligent adolescents." *Jour. Psychol.,* 4: 1937, 287–293.

(37) HOWARD, D. T. "A functional theory of the emotions." [In] *Feelings and Emotions: The Wittenberg Symposium.* (Clark University Press: Worcester, Mass., 1928) 454 pp. (pp. 140–149)

(38) KNIGHT, M. W. "A comparative study of the performance of feeble-minded and juvenile delinquents on the Arthur Performance scale and the Stanford-Binet tests of intelligence." *Jour. Juv. Research,* 18: 1934, 5–12.

(39) KÖHLER, W. *The Mentality of Apes.* (Harcourt, Brace: New York, 1925) 342 pp.

(40) KOLB, L. "The relation of intelligence to the etiology of drug addiction." *Amer. Jour. Psychiat.,* 5: 1925, 163–167.

(41) KOLB, L. "Types and characteristics of drug addicts." *Mental Hygiene,* 9: 1925, 300–313.

(42) LANE, H., and WITTY, P. A. "The mental ability of delinquent boys." *Jour. Juv. Research,* 19: 1935, 1–12.

(43) LANGFELD, H. S. "Judgments of facial expression and suggestion." *Psychol. Rev.,* 25: 1918, 488–494.

(44) LANGFELD, H. S. "The judgment of emotions from facial expressions." *Jour. Abnorm. Psychol.,* 13: 1918–1919, 172–184.

(45) MCCAULLEY, S. "One hundred non-conformed boys." *Psychol. Clin.,* 16: 1925, 141–166.

(46) MACKAYE, D. L. "The interrelation of emotion and intelligence." *Amer. Jour. Sociol.,* 34: 1928, 451–464.

(47) MILES, W. R. "Psychological effects of alcohol in man." [In] *Alcohol and Man.* H. Emerson, Ed. (Macmillan Co.: New York, 1932) 451 pp. (pp. 224–272)

(48) MINER, J. B. *Deficiency and Delinquency: An Interpretation of Mental Testing.* (Warwick and York: Baltimore, 1918) 355 pp.

(49) PESCOR, M. J. *A Statistical Analysis of the Clinical Records of Hospitalized Drug Addicts.* (United States Public Health Reports: Washington, D. C., 1938) 32 pp. (Supplement No. 143)

(50) PINTNER, R. *Intelligence Testing: Methods and Results.* (Holt & Co.: New York, 1923) 406 pp.

(51) SEGAL, H. L., BINSWANGER, H. F., and STROUSE, S. "Effect of emotion on basal metabolism." *Arch. Internat'l Med.,* 4: 1928, 834–842.

(52) SELLING, L. S., and STEIN, S. P. "Vocabulary and argot of delinquent boys." *Amer. Jour. Sociol.,* 39: 1934, 674–677.

(53) SIMMONS, R. McK. *A Study of a Group of Children of Exceptionally High Intelligence Quotient in Situations Partaking of the Nature of Suggestion.* New York, 1938. (Doctor's dissertation to be published in Teachers College, Columbia University, Contributions to Education)

(54) SKEELS, H. M., and FILLMORE, E. A. "Mental development of children from

underprivileged homes." *Ped. Sem. and Jour. Genet. Psychol.*, **50**: **1937, 427–439.**

(55) SLAWSON, J. *The Delinquent Boy: A Sociopsychological Study.* (Richard G. Badger: Boston, 1926) 477 pp.

(56) STOKE, S. M. *Occupational Groups and Child Development.* (Harvard University Press: Cambridge, 1927) 92 pp. (Harvard Monographs in Education, No. 8)

(57) STRATTON, G. M. "The function of emotion as shown particularly in excitement." *Psychol. Rev.*, 35: 1928, 351–366.

(58) SULLENGER, T. E. *Social Determinants in Juvenile Delinquency.* (John Wiley: New York, 1936) 412 pp.

(59) SULLIVAN, E. B. "Age, intelligence, and educational achievement of boys entering Whittier State School." *Jour. Delinq.*, 11: 1927, 23–38.

(60) TERMAN, L. M. *The Measurement of Intelligence.* (Houghton Mifflin Co.: Boston, 1916) 362 pp.

(61) TERMAN, L. M. *The Intelligence of School Children.* (Houghton Mifflin Co.: Boston. 1919) 317 pp.

(62) U. S. CHILDREN'S BUREAU. *Juvenile Court Statistics and Federal Juvenile Offenders, 1933.* (United States Department of Labor. Government Printing Office: Publ. 232. Washington, D. C., 1936) 119 pp.

(63) WALL, J. H. "A study of alcoholism in men." *Amer. Jour. Psychiat.*, 92: 1936, 1389–1401.

(64) WALL, J. H. "A study of alcoholism in women." *Amer. Jour. Psychiat.*, 93: 1937, 943–955.

(65) WASHBURN, M. F. "Emotion and thought: A motor theory of their relations." [In] *Feelings and Emotions: The Wittenberg Symposium.* (Clark University Press: Worcester, Mass., 1928) 454 pp. (pp. 104–115)

(66) WEBER, C. O. "Moral judgment in female delinquents." *Jour. Applied Psychol.*, 10: 1926, 89–91.

(67) WITTY, P. A. *A Study of Deviates in Versatility of Play Interests.* (Teachers College, Columbia University, Contributions to Education, No. 470: New York, 1931) 57 pp.

(68) WITTY, P. A., and LEHMAN, H. C. "Nervous instability and genius: Some conflicting opinions." *Jour. Abnorm. and Soc. Psychol.*, 24: 1930, 486–497.

(69) YODER, A. H. "The study of the boyhood of great men." *Ped. Sem.*, 3: 1894, 134–156.

CHAPTER XI

NEW EVIDENCE ON ENVIRONMENTAL INFLUENCE ON INTELLIGENCE

FLORENCE L. GOODENOUGH
Research Professor, Institute of Child Welfare
University of Minnesota
Minneapolis, Minnesota

I. INTRODUCTION

The word 'new' as used in the title of this chapter has reference to data appearing since the publication of the *Twenty-Seventh Yearbook* of this Society in 1928. It does not include the fairly large body of critical discussion and re-interpretation of the data presented in that volume that has continued to appear up to the present time. The point of view to be presented here and the manner of approach to the question as a whole differ somewhat from those of the earlier discussions, in that the concept of 'environmental influence' will be envisaged in a highly specific rather than a generalized manner and a sharp distinction will at all times be made between the term 'intelligence' (in the sense of generalized ability as manifested in life situations) and 'intelligence-test score,' which may or may not be a useful index to that ability. We shall not attempt to answer such questions as "What percent of individual differences in intelligence is due to heredity and what percent to environment?" but instead we shall ask "What effect, if any, is a specific environmental change likely to have upon the intellectual development of a particular group of children?" Realizing that many factors unrelated to true mental differences may affect the results of even the best intelligence tests available at the present time, and that statistical fallacies frequently operate in such a manner as to give a spurious appearance of relations that do not in fact exist, or, on the other hand, to obscure relations or mask changes that have truly occurred,[1] we shall attempt to approach the problem of environmental influences upon mental growth in a spirit of critical inquiry. For the time being, at least, we shall assume the rôle of the doubting Thomas demanding evidence that is as nearly crucial as it can be made.

[1] See Part I, Chapter XII of this Yearbook.

This general problem of environmental influence upon intelligence is most easily discussed by subdividing the evidence according as it pertains to (*a*) the effect of nursery-school training, (*b*) the effect of later school training, and (*c*) the effect of living in foster homes.

II. The Evidence on Problem I: The Effect of Nursery-School Training on the Intelligence of Children

That great variations exist among nursery schools in respect to the educational philosophy underlying their management, the competence of the teachers, and the character and home backgrounds of the children who attend them is a fact too obvious to require proof. In considering the possibility of modifying the course of mental growth by attendance at a nursery school, there is not, therefore, one fundamental question, " Do all nursery schools bring about such an improvement? " but rather there is a series of questions like the following: " Does any nursery school do so, and if this is the case, by what special procedures that differ from those found in other schools does it accomplish this result? Can all children be benefited intellectually by attendance at this school, and if not, how may we know which ones are most likely to improve under such a nursery-school regime? Is the mental acceleration so induced temporary or lasting? If permanent, is the gain confined to the increment of mental age achieved during the nursery-school period or is there an acceleration in rate of growth that persists after the nursery school has been left behind? " [1]

[1] Such data as have been reported seem to imply that the latter is the case, although it is hard to see how this could come about. Wellman (41), for example, states that after an average attendance of one and a half years, the mean gain in intelligence quotient for 45 children first tested at 2, 3, or 4 years was 15.6 IQ points. Since the distribution of ages is not given, it is impossible to say with assurance what the age at final testing was, but assuming it to be 5 years, this would be equivalent to a mental-age increase of about 9 months. Now if this mental-age increment were merely retained without further acceleration in rate of growth, it would be equivalent at the age of 16 to an increase of less than 5 IQ points. While this is worth having, it does not seem to correspond to the great differences reported by the same author in a later article (44) in which she compares the performances of preschool and non-preschool children on college-entrance tests at the age of 18 years. It is true that the use of percentiles in the later study greatly exaggerates the apparent magnitude of the differences between the groups and without more detailed information about the form of the distribution of cases from which the percentile norms were derived, it is impossible to transmute the percentiles into intelligence quotients or vice versa. However, the differences reported appear to be considerably greater than would correspond to 5 IQ points.

By far the most extensive of the studies that appear to show a positive effect of nursery-school training upon the intelligence of children are those published by Wellman and her colleagues at the University of Iowa. With a few exceptions, the reports from other nursery-school centers show no effect. Needless to say, this fact in no way demonstrates that the Iowa findings are incorrect, since it is entirely possible that, as Wellman so ably points out in an earlier chapter, the gross features of an environment provide us at best with only a very crude measure of its psychological significance for the individual. It is entirely possible that a system of preschool education has been developed at Iowa that brings results that others have so far failed to duplicate. However, it cannot be denied that the absence of measurable effect upon intelligence in the case of nursery schools under such able direction as those at the University of California, the University of Minnesota, the Winnetka public-school system, the University of Cincinnati, and others renders it essential that the experimental evidence be examined with more than usual care. The question is: "Do the apparent differences correspond to a genuine difference in the stimulating power of the schools in question or can they be traced to experimental or statistical errors on the part of one or another of the investigators in question?" In the one case we must inquire, "Is the apparent gain real or fallacious?" In the other case, "Is the apparent lack of improvement due to failure on the part of the investigators to tease out the significant relationships involved?" Only if the differences cannot be adequately accounted for on technical grounds need we concern ourselves with a search for explanatory principles in the idiosyncrasies of nursery-school management.

1. Early Studies

Except for the Iowa studies and those included in Part II of the present Yearbook, the number of investigations on the effect of nursery-school training is small.

Of the three appearing previous to the time covered by this survey, that by Woolley (50) appeared to show at least temporary effect, but since the children were not followed after leaving nursery school, its significance for later development is unknown. Neither can it be said with assurance whether or not the gain in test-performance may not have been brought about wholly or in part by specialized experience of a kind so similar to the test-items as to have the effect of unintentional coaching.

The second report, by Hildreth (7), shows that children entering the first grade after at least four months of nursery-school or kindergarten training tested, on the average, about 6 points higher than a similar group entering without previous school experience, but that at the end of 18 months the first group had registered a slight loss, the second a slight gain, with the result that the initial difference of 6 IQ points had been reduced to less than 2.

The third report, by Goodenough (5), shows a statistically insignificant difference between the changes in test standing of a group of children after a year of nursery-school training and that of a group of paired controls. A further study by Goodenough and Maurer will be found in Part II of this Yearbook.

2. The Study by Barrett and Koch

Except for the Iowa studies, I have been able to locate only two appearing since 1928 that purport to show an accelerating effect of nursery-school training upon child intelligence. In 1930, Barrett and Koch (3) reported findings from an orphanage project that, except for its brief duration and the smaller number of cases involved, is similar to the one set up in Iowa by Skeels, Wellman, Updegraff, and Williams (28). Seventeen pairs of young orphanage children were used as subjects. All members of the experimental group attended a single nursery school set up in one of the four orphanages from which the control subjects were drawn. The fact that not all the control children were inmates of the same institution may have some bearing on the results, but this seems unlikely because of the careful matching of the pairs on the basis of age and initial test-status. However, it would have been desirable to have more information on this point, since all cases in the experimental group were taken from the Chicago Orphan Asylum, whereas nearly all [1] the control-group cases were taken from three other institutions said to be similar to the first in general equipment and in the physical management of the children. No information is given about the racial or religious affiliations of the several orphanages or other differences in the social background of the children that might have a bearing upon growth potentialities.

The Merrill-Palmer Tests were used for both measurements, which were made at the beginning and the end of the first year of nursery-school training. Both groups showed some gain in standing at the time of the second test, but the experimental group gained much more than

[1] Exact number not stated.

the control group. Fifteen of the 17 cases gained more than the median gain of the controls and each individual member of the nursery-school group gained more than his paired control. A suggestion that the differential gain might perhaps be related to a greater improvement on the part of the nursery-school children in personality traits affecting the score is made by the authors, who report that the three children showing the greatest increase in test standing were judged by their teacher (who was not aware of the test results) to have improved most in personality.

Whether or not it is fair to assume that the difference in test score represents a genuine improvement in intelligence in the more narrow sense of the term — a difference that might be expected to persist into later years — cannot be said from the data at hand. It may be remarked, however, that the results of this and other experiments do offer support for the value of nursery-school training, regardless of whether the reported increase in test performance be due to a gain in the ability to think in abstract terms or to an improvement in general emotional and social adjustment.

3. The Study by Starkweather and Roberts

The second study showing improvement in test standing as an apparent result of nursery-school training is reported by Starkweather and Roberts (33) in Part II of this Yearbook. Since the authors state that their approach to the problem and their method of handling data are taken over without essential modification from those used at Iowa, we shall defer the discussion of this study until after the Iowa material has been considered.

4. The Studies of Iowa City Preschool Children

The Iowa evidence may be grouped under two heads: (*a*) changes in IQ of Iowa City children attending the University of Iowa preschools, analyzed in various ways and compared with changes in IQ of children not attending preschool, (*b*) a similar comparison between children attending a nursery school set up in the Iowa Soldiers Home, an orphanage and child-placing agency at Davenport, Iowa.

The evidence on the first of these heads is set forth in a series of articles dating from 1932 to the present time (41, 42, 43, 44, 47). For the University of Iowa group, the weight of evidence as to the effect of preschool attendance rests chiefly upon the analysis of changes within the preschool group. The number of control children studied is for the most part small and not always comparable with respect to age, level of intelligence, and test experience.

For example, in the fourth of the publications listed the percentile ranks on college-entrance examinations of children who had in early childhood attended the University of Iowa preschools are compared with those whose elementary-school experience was similar but who had not attended a preschool. Both groups had received Binet tests in childhood, but, according to Wellman's 1932 article (41), the preschool children had been tested twice a year while attending preschool and annually thereafter, whereas those entering the elementary school without preschool experience were tested but once, presumably at the time of entrance.

Although the college-entrance examination showed a slight superiority of the preschool group in terms of percentile rank, even when matched for childhood IQ, this fact might easily be accounted for by greater experience in testing on the part of the preschool children. The data are complicated further by the practice of carrying out various arithmetical computations (adding, subtracting, averaging) with the unequally spaced percentile ranks.

The indefensibility of this procedure, which is characteristic of most of the studies from the Iowa laboratory, has previously been pointed out by Simpson (23) and by McNemar (15). Although, as Wellman points out in her reply (48), Simpson's estimation of the extent of the error is greatly in excess of the facts, the discrepancies induced by this practice are still large enough to produce many irregularities in the numerical results.

Some data comparing preschool and non-preschool groups are also given in Wellman's third paper, but as the groups are unequally matched in respect to age and testing experience previous to the tests given at the research station and labelled "first test," and since no information is given with regard to home background or other selective factors that might be related to growth potentialities, it seems better to ignore this line of evidence and turn to that derived from repeated testing of the preschool children alone.

Considering first the general direction of the change in IQ with repeated testing, it is shown that the trend for the preschool group as a whole is upward. This is in accordance with the findings of other investigators reported in Part II of this Yearbook.[1] However, those who have compared the changes in IQ of nursery-school children with those made by carefully matched controls have usually found that the lat-

[1] See especially the reports by Goodenough and Maurer, Voas, Bayley, and Frandsen.

ter also gain to approximately an equal degree.[1] It is true that the average gain reported for the Iowa group is somewhat greater than that found by most other investigators. Whether the differences should be ascribed to a difference in the stimulating character of the schools or to different conditions of testing cannot be said from the data at hand, but because of the possible effects of unconscious bias,[2] it would seem desirable to check these findings by tests given by someone unacquainted with the children or their educational history.

In addition to the general comparisons based on the group as a whole, the material has been subjected by the investigators to many types of secondary analysis, according to duration of nursery-school enrollment, number of days' attendance, gains from fall to spring compared with those from spring to fall, and the like. Data are also presented on changes in specially selected individuals. Typically, the results are presented separately for groups of varying initial test-standing, with the result that regression due to errors of measurement — always higher at the preschool ages than at later ages — causes the children originally testing high to rank somewhat lower on later tests, whereas those originally testing low appear to make very marked gains.[3] Since the various factors commonly subsumed under the heading of 'practice effect' have almost universally been found to bring about an upward shift in the mean of a second test when compared with the first, the regressive 'gains' made by the low-testing group are commonly greater than the 'losses' by the children initially testing high. The Iowa authors have developed an ingenious theory to account for this differential change, which, in spite of the extensive literature on the subject in the field of mathematical statistics, they continue to look upon as 'real.' The theory, as set forth in a number of the Iowa papers, appears to be that in any situation there is a critical point toward which the stimulating value of the environment tends to

[1] Possible explanations for this rather general tendency have been discussed in Part I, Chapter XII, of this Yearbook.

[2] This statement is not to be construed as a suggestion that the Iowa results were *intentionally* affected by the attitude or procedures of the examiners. Science, however, is replete with instances in which the enthusiasm of individuals who believed themselves to have made a new scientific discovery has caused them to overlook some of the apparently minor factors in experimental technique that were later shown to have been of major importance in determining the results. The story of the famous horse, *Clever Hans,* is a classical example. It is pertinent, therefore, to examine the Iowa results with care to ascertain what special precautions, if any, were taken to guard against such possibility.

[3] See Part I, Chapter XII.

impel the intellectual development of the children. Those above this point lose for lack of further stimulation; those below it, gain. The theory is an interesting one, but before it can be regarded as worthy of serious practical attention, it must be shown that the amount of the differential shift is greater than that necessitated by regression due to errors of measurement. Such a computation can easily be made where the basic data are available; it is surprising that thus far it has not been done.

Several of the Iowa investigations have shown that the gains in IQ of nursery-school children are greatest between the period from fall to spring, whereas during the summer months the tendency is toward a small loss. Two things are somewhat puzzling in these reports. In the first study in which this problem is considered (42), the comparison is made for nursery-school children only. In that study, as in many others, the problem is complicated by transmuting the IQ's into percentiles before making the comparison. The reason advanced for this transmutation is the drop in mean IQ with age due to incorrect standardization of the 1916 Stanford-Binet, but since this drop does not occur until a considerably later age than that of nursery-school children (Wellman's 1932 study shows no consistent change until after the age of eight years), the shift to the percentile basis is not only undesirable, but also unnecessary. Be that as it may, it appears from Figure 3 of the study (42) that, with each year of nursery-school attendance, the *loss* during the ensuing summer steadily increases. Just what this signifies is conjectural; it may be a spurious result of the averaging of percentiles of unequal significance, but this cannot be said with certainty since it is impossible to unscramble the figures and reduce them to their original form. One would also like to know by what statistical process the totals shown in Wellman's Figure 3 were obtained.

In the 1934 article (43), data are given on fall-to-spring changes in IQ for 34 pairs of nursery-school and non-nursery-school children of closely similar age and initial IQ (as taken in the fall). The preschool children gained an average of 7 points; the non-preschool children lost, on the average, 3.9 points. In computing the critical ratios for these results, one is justified in criticizing Wellman's line of reasoning. The question is whether or not the 3.9 IQ points lost, on the average, by the children of the control group, are or are not to be regarded as a real loss. Since the critical ratio of the difference between the fall and spring tests is only 0.97, one would be quite justified in regarding it as a chance fluctuation from a point of zero change. If this interpretation is adopted, however, then the question is narrowed to a simple consideration of the reliability of the gains made by the nursery-school group, without any reference whatever to the changes in the control group, which are then assumed to have no real existence. The critical ratio given by Wellman for the difference between fall and spring IQ's for the nursery-school group is 2.07. In her discussion, however, she stresses the critical ratio obtained from the difference between the spring tests of the two groups (7.0 + 3.9 + 0.1, the last being the initial difference in favor of the preschool group, which, by this method, is now implicitly attributed to nursery-school training). But this procedure carries with it the necessary assumption that the loss of 3.9 IQ points by the control group was just as 'real' as the gain of 7.0 by the preschool group, and calls equally for explanation. One must then wonder what went wrong in the homes of the control children that caused them to become more backward.[1]

Since other types of analysis in these studies leave room for considerable doubt as to their real significance, it is sufficient to say that, while there is no reasonable doubt that a numerical increase in mean IQ has truly been shown by children attending the University of Iowa nursery school, it is unsafe to assume that this indicates a real gain in intelligence until at least the more serious of the technical errors in the collection and treatment of the data have been recognized and satisfactorily met.

5. Wellman's Study of the Merrill-Palmer Scale

A study of the changes in performance on the Merrill-Palmer Scale of Performance Tests (47) can also be dismissed with brief comment. This test is not well suited to research purposes because of the difficulties involved in scoring.

As Stutsman correctly points out, the relationship of chronological age to the standard deviation of mental age necessary for the valid use of the IQ is not fulfilled for this test, so that the expressing of test results in terms of IQ is not valid. The errors involved in the use of the IQ are of considerable magnitude, especially for children who diverge markedly from the average in ability. For example, according to Stutsman, " at +2.5 S.D. the IQ varies from 122 to 165; at +2.0 S.D., from 119 to 154; and at +1.5 S.D., from 114 to 141. It is apparent that with this amount of variation, such an index has little significance." Stutsman (35, p. 106) suggests that in place of the intelli-

[1] Cf. the mean loss of 4.6 points in the orphanage experiment (28) described farther on.

gence quotient, percentile ranks or sigma indices should be used. As already noted, the former do not lend themselves readily to statistical treatment, and the sigma norms given by Stutsman are expressed in such coarse units (1/2 S.D.) that interpolation is necessary for careful research work.

The majority of children studied by Wellman were from the University of Iowa nursery schools, where the level of intelligence is high, although a smaller number from other institutions was also included. Since, according to Stutsman (35, p. 237) children at 2.5 S.D. above the mean reach the 'ceiling' of this test at 49 months, it is evident that the scale is poorly suited to measuring the ability of very bright children much over the age of three and a half — a statement borne out by our experience with the test at Minnesota.

Wellman's subjects ranged in age from 20 to 62 months. Although she reports that children attending preschool gained more on this test from fall to spring than from spring to fall, and also finds a greater gain for children who attended the nursery school than for those who did not, she also notes that there were irregularities in the results. This is not surprising in view of the irregularities in the test, particularly when, as is true in many of the computations, intelligence quotients from children of different ages are grouped together in a single table. When an IQ of 114 at one age has the same meaning as one of 141 at another age, almost anything can happen.

It is unfortunate that a test having so many excellent qualities does not lend itself better to research purposes. When, as in the experiment by Kawin and Hoefer to be described later on, the only question involved is a comparison between changes in groups closely matched, case for case, for age at initial and final testings, much of the difficulty is ironed out, but this condition seems not to have been fulfilled in Wellman's study.

6. The Iowa Study of Orphanage Children

The report of the orphanage project by Skeels, Updegraff, Wellman, and Williams (28) has been widely quoted as affording evidence that lack of environmental stimulation may reduce children originally normal to the level of mental deficiency. The evidence for this oft-repeated statement is obtained from a comparison of the changes in IQ of children living in an orphanage under conditions dramatically described as extremely lacking in psychological stimulation of all kinds — little or no opportunity for conversation with adults, and practically no toys or supervised play.

In this institution a model preschool was established. For every child admitted to this preschool there was a control child who remained in the restricted environment of the orphanage. Pairing was done on the basis of chronological age, mental age, sex, nutritional status, and length of residence in the orphanage. Owing, however, to the frequent removal of children from both groups for placement in foster homes, which necessitated their replacement in the study by other similar cases, adequate matching was hard to maintain — as a matter of fact, as the experiment continued, it so worked out that the 46 children who were at one time or another enrolled in the preschool group averaged 4.5 IQ points higher on the initial test than the 44 children who served as controls.

The average amount of change in the measured IQ of these children is surprisingly small in view of the rather startling pronouncements made in regard to it. According to the authors' table, attendance at the orphanage preschool for periods ranging from 1 to 199 days resulted in a mean loss of 0.7 IQ points, during which time the control children gained 0.2 points. The authors account for the absence of change during this time on the basis of the great difficulty experienced by the teachers in adjusting the children to the new regime. After attendance for 200 to 399 days, however, the preschool children showed a mean gain of 3.7 points; the controls, a loss of 1.2 points. After 400 or more days of attendance, the preschool children had gained an average of 4.6 IQ points, whereas the controls had lost an equal amount. This hardly seems sufficient warrant for the statement on page 57 (which is in substance repeated elsewhere in the monograph and in numerous popular articles) that "another conclusion which seems to emerge [from this study] is that children of average ability may be made feebleminded." True, the reference here is not to the *average* child in the control group (since a drop of 4.6 IQ points is obviously a small fraction of the span from normality to mental deficiency), but to changes occurring in a small number of individual cases where losses appeared as great as 25 to 43 IQ points. Whether these apparent losses represent true retrogression in intelligence or failure of the examiners to secure adequate rapport on the second test[1] is a moot

[1] The tests, it is to be noted, were given by a large number of different examiners. It is highly unlikely that all were equally skilled in securing coöperation from children as difficult to handle as those of the control group are said to have been. In this connection it is also to be noted that one effect of nursery-school training about which practically all who have had experience in the matter are agreed is that it makes children easier to test.

question, but from an examination of the original data that were supplied by Wellman, McNemar (15) ascertained two points that are of considerable importance in this connection. First, of the 7 control-group children whose marked losses occasioned the conclusion that "children of average ability may be made feebleminded," *5 were under 24 months of age at the time the diagnosis of "average ability" was made and the three whose losses were greatest (43, 37, and 24 IQ points) were under 20 months at the time of the initial test.* McNemar also points out that there were 7 control group children and only 3 of the preschool group who showed *gains* from initial to final testing of more than 14 points.

Other comments might be made on this study, which covers many other aspects of development as well as that revealed by the Binet tests, but it seems unnecessary because the remainder of the report is similarly open to question.

We may summarize by saying that the authors have made a good case for the advantages provided by the orphanage preschool in increasing the happiness of the children who attended it and presumably in improving their general behavioral adjustment. As to the imputed changes in intelligence, the data presented are insufficient to warrant any conclusion whatever, especially in view of the questionable statistical procedures employed. (These are not reviewed here since they have previously been pointed out by McNemar.)

7. The Study by Starkweather and Roberts

Let us now consider the study by Starkweather and Roberts (33). The reader will note that their statistical procedures are taken over directly from those used by Wellman and are consequently subject to the same criticisms. Examination of their Table 5 apparently shows, among other things, that it is very damaging to the mentality of a highly superior child to send him to the Merrill-Palmer Nursery School! The *loss* in IQ for these children following attendance (and therefore presumably resulting from this attendance) is on the average 10.3 IQ points (critical ratio 4.25). It should be unnecessary to explain that this loss, like the contrasting gains at the low end of the scale, is an effect of regression. The Iowa practice of transforming IQ's into percentiles and then averaging the percentiles is also copied faithfully, both for the Binet and the Merrill-Palmer tests. There is no control group.

8. The Study by Kawin and Hoefer

Now for the contrary considerations, which are represented by a number of studies. Most of them are to be found in Part II of this Yearbook and will be dealt with presently.

Of the studies not in the Yearbook, a report by Kawin and Hoefer (8), although based upon a small number of cases, is of particular interest because of the meticulous care with which all possible variables were controlled. The subjects were 22 closely matched pairs of children tested at the beginning and the end of the experimental period. During this time, one member of each pair attended a nursery school; the other did not. In all measured respects, the home conditions of the children in each group were essentially similar and their mean chronological and mental ages at the time of the initial test differed by only a small fraction of a month. The average attendance record was 78 percent of the total possible time during the school year covered by the experiment. No nursery-school child was included whose attendance record fell below 50 percent. The Merrill-Palmer tests were used for both the initial and final examinations.[1] An important feature of this study is the fact that all examinations were given by an examiner who had no other contact with the children and thus was no more familiar to the nursery-school children than to the control group.

Both groups ranked higher (in comparison to age) in the spring than they had in the fall. This is an almost universal outcome with the Merrill-Palmer tests, since the effect of practice upon performance seems to be unusually high in them. The gain, however, was practically identical for the two groups. The exceptional regularity of the findings shown by a mean gain of 11.3 mental months for the nursery-school children and of 11.4 months for the controls after a mean interval between testings of slightly less than 7 months for each group is in all probability the result of the careful attention to experimental technique throughout the duration of the study.

9. Studies Included in Part II of this Yearbook

The investigations by Bird, Frandsen, Lamson, Olson, and Anderson in Part II of this Yearbook need little comment. All are based on

[1] Since at the time of the final test the oldest child in the group was only 49 months of age and the majority were under 3½ years, the age factor previously mentioned for these tests was not a serious difficulty.

one year of nursery-school experience, and none indicates any reliable effect of this experience upon intelligence-test performance. Although the number of cases in each is not large, the total number represented by the five studies is of respectable size, especially when taken in conjunction with the more extensive studies by Bayley, Goodenough and Maurer, and Voas, which show similar results. Whether the differences in the findings from the various sources are to be attributed to differences in the stimulating value of the schools or to differences in testing procedures and the statistical treatment of data will be left to the reader to decide.

10. The Study by Peterson

One more study need be mentioned here, which, curiously enough, is a product of the Iowa laboratory. This is a study by Peterson (18) in which a group of children who had previously attended the University of Iowa Nursery School is compared with a group who had had no nursery-school training at the time of entrance to the University Elementary School. The measurements included a large number of physical, mental, and social traits. In general, the differences were small and unreliable. The point of chief interest, however, is not so much the study itself as an abstract of it by Wellman (45), which runs in part as follows: " The nursery-school group had gained in IQ from the time of entrance to preschool to enrollment in junior primary. Scores in reading readiness, vocabulary, and information indicated little difference between the groups." Reference to the original study, however, shows that the difference between the mean IQ of the nursery-school group and that of the non-nursery-school group at the time of entrance to the junior-primary school was only 3.6 IQ points. This difference, which is in favor of the nursery-school group, is equivalent to a critical ratio of 0.675. At the end of the school year, the difference had dropped to 2.6 IQ points, or a critical ratio of 0.49. On the three tests said to show " little difference between the groups," the differences were all in favor of the group *without* nursery-school experience. For these tests the critical ratios were as follows: on the Probst test of general information, 1.49; on the Smith-Williams vocabulary test, 1.28 at entrance and 1.75 at the end of the school year; on the test of reading readiness, 0.676. In view of these figures, the statement quoted from Wellman's abstract is not easy to understand.

11. Summary of the Effect of Nursery-School Training

Tentatively, at least, the questions propounded at the beginning of this section may be answered as follows:

1. It is unsafe to assume that attendance at any unspecified nursery school is likely to bring about improvement in the mental ability of the average child, for in most investigations of this matter, no evidence whatever of such an effect has been found. It appears, therefore, that the ability to bring about intellectual improvement by this means is at best restricted to a few schools. The precise nature of the differences in the educational regimes of the schools that do, and of those that do not, claim to achieve these results has thus far not been made clear.
2. An analysis of the experimental studies dealing with this problem reveals many possible sources of error in the studies purporting to show a positive effect of nursery-school training. Generally speaking, these studies have failed to maintain adequate control of basic variables. Their statistical techniques are also frequently of questionable validity, and in a number of instances, are certainly erroneous. A similar and more complete analysis than that given here has been published by McNemar (15). Unless McNemar's criticisms and those of others who have called attention to these matters can be satisfactorily answered, the remaining questions in our list require no further comment, since their appropriateness is contingent upon the reply to the first.

Critical readers will agree that the Iowa claims are far less substantial than has seemed to those only casually acquainted with them. Many will feel that there is still room for a more crucial test.

12. The Crucial Test

In all science, the crucial test of any new claim is held to be that of *verification by independent investigators under similar conditions.* " The proof of the pudding is in the eating " is a popular expression of a similar idea. No competent man of science should object to having the principle applied. As Wellman has very properly stated (Part I, Chapter XIV of this Yearbook), the psychological aspects of a given environment may be very different from its physical aspects, and the fact (evident from a consideration of the studies included under Part II) that not all nursery schools are successful in improving the intellectual level of the children who attend them is no proof that others may not do so. But if this be granted, then it becomes incumbent upon those who claim such accomplishment to accept for training other children whose initial ability has been established by independent investigators who can later test the results of the training. There are thousands of anxious parents throughout the country whose children show early signs of mental retardation who would gladly pay any reasonable cost in order to insure their normal developmental progress. There are thousands of others who would do likewise for the sake of causing their " originally average " children to " break over into the genius groups " (41).

I propose, therefore, that one of the two following plans be adopted by the members of the Society:

Plan A: (1) That an attempt be made to secure funds for an investigation in which a special group of dependent children is selected by an outside committee of competent scientists according to specifications laid down by the persons claiming the ability to bring about improvement in intellectual status; (2) that these children be placed in boarding homes selected by the persons to be responsible for their development; (3) that they would attend nursery schools under the direction of the same group; (4) that a second, comparable group of non-dependent children who are being reared in their own homes should be used as a control; (5) that at annual intervals thereafter, the children of both groups be examined by a second committee of scientists ignorant of the group to which the individual children belong.

While the cost of such an investigation carried out over a period of years would be considerable, it would not be beyond reasonable hope of attainment if a concerted effort were made to secure the funds. All details of the investigation should be carefully planned in advance by the three groups coöperating in the study, but it should be understood that the group responsible for the training program would be left entirely free to employ any methods of training that they saw fit to use and that those responsible for making the examinations would likewise be free to use any methods of appraisal that seemed desirable, provided, of course, that the same methods were used for both the experimental and the control group.

Plan B: While there is no question that the controlled experiment described as Plan A would be far preferable to the more haphazard plan of confining such a study to children whose parents were willing to bear the necessary expense, there is little question that many such cases could be found. Plan B would therefore be that of recommending parents of backward children or those whose children are not so superior as is desired to enter them in the nursery schools in question in order that their ability might be improved. Even a small number of test cases, *carefully studied by independent and qualified observers,* would afford valuable information as to the probable validity of the claims that have been put forth, though the test would be less crucial than that first proposed because of the greater possibility of uncontrolled variables affecting the results and the large error of measurement involved when small children are used as subjects. However, the plan has the obvious merit of requiring no special funds, since the expenses would be borne by the parents themselves. It is therefore a feasible one.

III. The Evidence on Problem II: The Effect of Later School Training on the Intelligence of Children

1. The Question of Differential Practice

Since our compulsory education laws demand that all children shall attend some kind of school after the age of 6 or 7 years, evidence on the effect of later school experience is for the most part restricted to a comparison of children who have attended schools of differing types. Such comparisons are open to many hazards of selection, both with respect to the children themselves and with respect to special factors in school practice that may affect the reliability and validity of the test results. That there is a practice effect resulting from past experience in taking tests has been repeatedly shown; hence, granting equal ability, the children in a school where tests are frequently given will commonly earn higher scores than those who are tested for the first time. This is probably responsible, at least in part, for the findings of Wellman (44), who reports a higher mean standing on college-entrance tests for students who had attended preschool previous to enrolling in the University of Iowa Elementary School or High School than for those who had the latter training but did not attend the University preschool, since the former had had a far greater amount of test-experience than the latter.

2. The Question of Differential Selection: Maller's Study

That different school populations may truly differ in ability because of various social and cultural factors influencing the selection of children who attend them has been so often stressed that it is unnecessary to elaborate the point. Private-school children commonly test higher than public-school children; children from schools located in the better residential districts are superior to those from schools in the slums, and so forth. Maller (16), for example, in a study of all the fifth-grade children in 273 health areas in New York City (total cases 100,153) by means of the National Intelligence Test and the Pintner Rapid Survey Test [1] found almost unbelievably large differences in the mean IQ of children from different areas of the city. In the lowest area the mean IQ was only 74; in the two highest areas the mean was 118. The stand-

[1] Both tests were given to all children and the average of the results used as the score for each child.

ard deviation of the means of the 273 areas was more than seven times as high as would have been expected had each area been a random sample of the total population tested.

3. The Iowa Studies

The Iowa studies on this topic have been summarized by Wellman and Skeels in the second part of this Yearbook. Since many of the original studies discussed in this review have been commented on by McNemar, it is unnecessary for me to make more than a few additional comments. The greater gains by the children originally testing low, together with the tendency of those with initially high standing either to lose or, at most, to make very slight gain, are probably attributable to regression. It should also be noted that the effect upon test scores of attendance at the University Elementary School seems to have been transitory. By the end of four years the mean IQ for the entire group was but slightly (less than 2 IQ points) higher than it had been at the time of entrance.

4. R. L. Thorndike's Study

Except for the Iowa series, few studies on the effect of later school environment upon change in intelligence have appeared, and those that have been reported are commonly hard to interpret because of (*a*) differences in testing practice in the different schools studied, (*b*) inadequate information about factors in the home backgrounds of the children that might be related to different potentialities of growth, and (*c*) the possibility of systematic errors of measurement that vary from age to age. Moreover, such differences as have been found do not always seem readily explainable on the basis of observable differences in the curricula of the schools compared. In Part II of this Yearbook, R. L. Thorndike (40) was unable to account for the differences in the mean amount of gain on retest shown by the pupils of his School B in comparison with the two other schools studied by him. A number of hypotheses designed to throw light on this question were tested, but none of them was verified. It should be noted, however, that a much greater number of examiners took part in the testing program in this school than in either of the others. In view of the apparent similarity of the schools in all matters that might be assumed to bring about a difference in rate of growth, further evidence is needed to demonstrate whether the differential gain should be attributed to (*a*) chance fluctuations of sampling, (*b*) differences in test administration when many different examiners were employed, or (*c*) a true difference in the stimulating value of the different schools.

5. The Study by Reymert and Hinton

The study by Reymert and Hinton (19) will be considered in more detail in connection with the studies of foster children discussed shortly. For the moment it suffices to note that the intelligence quotients of the Mooseheart children, who were past the age of 6 years at the time of entrance, did not increase as the result of attendance at the Mooseheart schools. The significance of this fact cannot be determined with assurance, since it depends upon the degree of contrast between the schools attended previous to entering Mooseheart (obviously a decidedly variable factor because the children came from many parts of the country) and those in the institution.

6. Klineberg's Study of Negro Children

Klineberg's study (10) of the performance on mental tests of Negro children with varying lengths of residence in New York City has been reviewed by Hollingworth and Witty (Part I, Chapter IX) in their discussion of racial differences.[1] However, it may be well to note here that Klineberg's own interpretations of his findings appear not to be wholly in acccordance with those attributed to him by some of his commentators. As a matter of fact, Klineberg's statements in different parts of the monograph are not always consistent with each other. For example, in commenting on the results of the performance tests — which showed no difference between groups of longer and shorter residence in contrast to the Stanford-Binet results wherein the northern-born exceeded the southern-born, though on the average by a smaller amount than was shown when group tests such as the Otis and the N.I.T. were used as the measure of intelligence[2] — Klineberg says (p. 52):

[1] The reader may also consult a review of Klineberg presented by Stoddard and Wellman in Chapter XIV. — *Editor.*

[2] At the beginning of his chapter dealing with the results of the Stanford-Binet test, Klineberg states that this test differentiated the groups far more clearly than did the group tests. This statement seems to be based on a hasty examination of only a small part of the evidence. Three studies using this test are reported. Only the one embracing the smallest number of cases shows large differences between residential extremes; in the other two studies with larger samples the differences are small. On the average, the Stanford-Binet IQ's are much less clearly related to length of residence than are those obtained by the group tests. Of this, Klineberg seems to be cognizant when drawing the conclusion quoted above. The earlier statement is perhaps an oversight.

> When taken in conjunction with the results reported in preceding chapters, they [findings on performance tests] justify the inference that those tests in which school training enters are very much more affected by a change of environment than those in which the effect of schooling is reduced to a minimum. The environment probably has some effect on these latter, also, but it is not nearly so marked nor so immediate. One is tempted also to draw the conclusion that the results with performance tests show that there is merely a 'linguistic' and not a *real*[1] intellectual difference between the northern- and the southern-born groups and that the differences which appear are entirely a function of the environmental effect upon the particular test used. The writer inclines to that view, but does not feel that the material presented here justifies it entirely, since the exact relation between the abilities measured by linguistic tests and those measured by performance tests is still in doubt. (p. 52)[2]

7. The Study by Kephart

The chapter by Kephart (9) on the effect of specialized training within an institution upon the IQ of defective boys is interesting and suggestive. It would be more convincing if more complete data were presented on the exact nature of the training program, since it has long been known that both direct coaching (on the specific items of the test used) and indirect coaching (on items similar to those of the test) will affect test standing, but this improvement is by the majority of persons not regarded as synonymous with a genuine change in mental level. The example of training (dealing with the ability to see absurdities) given in Kephart's study strongly suggests that indirect coaching, at least, was involved. The fact that children whose mental defect was known to have an organic base did not profit by such training is quite in line with other findings in the field of abnormal psychology.

A further comment might be made in reference to the age of these subjects. Generally speaking, it has been shown that the chances of modifying the course of a developmental process are inversely related to the maturity level that has been reached at the time the measure is put into effect. To this extent, those who claim an increase in intelligence as a result of training at the preschool ages are in a somewhat better case, for a considerable amount of mental development is still to be accomplished after that age. But Kephart's subjects ranged in age from 15 to 18 years at the beginning of the experiment. One may then fairly raise the question whether the change in test standing is

---

[1] Italics are Klineberg's.

[2] Quoted by permission of the Columbia University Press.

indicative of further ' mental growth,' in the ordinary sense of the word, or whether it might not better be ascribed to the acquisition of specialized skills and knowledge of a kind similar to those included in the test. No one, as far as I am aware, has taken the position that high-grade morons are incapable of learning. One would like to see Kephart's experiment duplicated in other institutions for mentally defective children, like the institution at Vineland where the training program is known to be of exceptionally high caliber.

8. Seagoe's Study of Transients

The effect of frequent change of schools was studied by Seagoe (20), who gave group intelligence tests and educational achievement tests to all children who moved into a small school district in California during the school year. More than a third of these children left the district again before the end of the year. The average length of stay for the entire group was about three months, the average number of different schools previously attended was 3.7. For the most part the data are not reported in sufficient detail to warrant much generalization. Taken at their face value, they appear to indicate that the group as a whole was of normal intelligence (IQ in Grades IV and V, 105.0; in Grades VI to VIII, 97.5), but there was fairly marked retardation in school achievement as measured by standard tests and the percentage retarded in grade placement was more than twice as great as for the district as a whole. The number of children in the transient group was 93 out of a total school population of 414. The difference in the obtained IQ's for the upper and lower grades might be explained on any one of a number of hypotheses, none of which can be verified from the data at hand, but since different tests were used at the two levels and the number of cases is small, it may be best to disregard it. That the frequent shift in schools had a more deleterious effect upon educational standing than upon intelligence-test scores seems fairly well established, but this means very little, especially since it might plausibly be argued that the wider experience gained by frequent moving from place to place may be favorable to intellectual development.

9. Studies of Isolated Groups

The studies of canal-boat children by Gordon (6) and of mountain children by Sherman and Key (22) have frequently been cited as evidences of the effect of educational and cultural deprivation upon intelligence. The basis for this conclusion is: first, the low rating on standardized tests of the group as a whole; and second, the fact that with advancing age the mean IQ's earned by the children steadily decreased. In the case of the mountain children, the fact that the test performance of the children varied directly with the degree of isolation and cultural backwardness of the several communities is also regarded by the authors as evidence that the retardation was due to deprivation and not to genetic factors. However, from the more complete presentation of the material by Sherman and Henry (21), it is evident that another interpretation is possible. First, the extreme cultural and social isolation of the groups studied is evidently not a matter of geographic necessity. In describing the most backward of the mountain communities studied, the authors state " Colvin Hollow is close to a center of American civilization . . . less than eight miles from a hard-surfaced road. Automobiles have driven within three miles of it . . . within three miles is a large mountain summer resort " (p. 1). Under these conditions one has a right to suspect that the cultural and educational backwardness so vividly described is a result, rather than a cause, of mental backwardness of the people. As a matter of fact, the dysgenic character of the migratory movements by which these little mountain communities came into existence is pointed out very explicitly by the authors in the first chapter of the book, although they appear to have lost track of these matters later on. For example, we find the following:

> The fertile valleys . . . were first settled about the middle of the eighteenth century. . . . Many of the migrating families . . . probably the poorest equipped and the least determined — stopped by the way and settled alongside the Germans. They appropriated what land their predecessors had left — naturally the least fertile and desirable. . . . There was soon a population pressure. The more ambitious resumed the journey westward. The others were forced upward [into the mountain valleys]. (pp. 10–11)

Again,

> The crucial period came, it seems, about the beginning of the present century. The more energetic began to desert the log huts and migrate to other parts of the state. A few were able to establish themselves and quickly forgot their mountain associations. Many failed and returned. (p. 14)

And again,

> Marriage of relatives from first cousins to distant relatives is common. This has led Colvin Hollow to the practical loss of the family name. Almost everyone is named Colvin. (p. 15) [1]

Given two centuries of social anemia, during which time all the ablest members of the group have been continuously drained away, leaving only the intellectual and volitional weaklings to interbreed and reproduce their kind, need we seek further for an explanation of the state of educational backwardness and intellectual degeneracy found? Lack of schooling? But our pioneering ancestors did not find schools ready built in the wilderness. They made schools, and it did not require two centuries of residence for them to do so. Accordingly, I find it hard to accept the idea that the low IQ's of the mountain children are to be explained solely on the basis of educational deprivation. One is forced to ask: Why were they so deprived?

10. Discussion

Because of the many uncontrolled factors in all studies of the effect of later school attendance upon child intelligence, it does not appear that even tentative conclusions are warranted. It is practically impossible to find in the civilized world of today any large number of children for whom equality of genetic potentialities and inequality of educational opportunity can fairly be assumed.

> Choose a community with an educational system that is definitely below average in its measurable features and you will almost inevitably find that the intellectual level of the adults who are responsible for the character of the schools is low; choose a community of superior adults and you find in it superior schools. One of the first things that every intelligent group of pioneers in this country did was to establish schools for its children. It was not by chance that Harvard College was founded only sixteen years after the landing of the Pilgrims and that the maintenance of both elementary and secondary schools was required by Massachusetts law as early as 1647. Schools are man-made institutions and their existence and quality reflect the characteristics of those who gave them being.

This is not to say that all schools are equally good or that all who plan them have been equally well guided. It would, indeed, be amazing

---

[1] These three quotations are made by permission of Thomas Y. Crowell and Co., publishers, from *Hollow Folk,* by Mandel Sherman and Thomas R. Henry.

if the correlation between true excellence of schools and true ability of those who are responsible for them were perfect. But the relationship is certainly high enough to constitute a serious difficulty in the way of those who attempt to investigate the question of the possible facilitating effect of later schooling upon intellectual growth. In our studies of nursery-school training the problem is simplified to some extent by the possibility of comparing extremes, those who have had no such training with those who have had varying amounts. Moreover, since there is not uniform agreement, even among intelligent parents as to the value of nursery-school training for their children, the problems of selection, while serious, are nevertheless much simpler than they become later on when the comparison must be made between children all of whom attend some kind of school. If we grant the possibility that school training may affect intelligence, we must also grant the probability that any school is likely to be better than none at all. The differences between the stimulating value of any two schools chosen at random is thus likely to be far less marked, and the measuring instruments used for determining the effect upon child intelligence must be correspondingly much more delicate, if such effects are to be dependably ascertained.

Generally speaking, even the attempts to demonstrate the differential effects of different kinds of school practice upon child achievement have been disappointingly meager when suitable controls have been employed.[1] Several explanations are possible for these findings: (1) The matters chosen for comparison may not be the things that are really important in determining progress; (2) the instruments of investigation may have been too crude or poorly chosen for the purpose in hand; or (3) differences inherent in the various individuals may so greatly overweigh in importance such differences in opportunity as are commonly found that the latter are almost, if not completely, lost by comparison. Probably each of these factors has had its part in confusing the evidence; thus far no one seems to have been able to determine which of the tangled threads is the one that binds the knot. If that be granted, we should not be surprised at the conflicting results of the attempts to demonstrate the effect of school training upon the more remote and less clearly defined characteristic that we call 'intelligence.'

---

[1] See several contributions in the *Twenty-Seventh Yearbook* of this Society, Part II, 1928 (17).

IV. The Evidence on Problem III: Children Reared in Foster Homes

1. Leahy's Study of Foster-Children

The discrepancies between the findings and conclusions of the studies by Freeman and those by Burks on the intelligence of foster children, both of which appeared in the *Twenty-Seventh Yearbook* of this Society, stimulated Leahy (12) to undertake a similar investigation under conditions that offered somewhat better opportunities for experimental control. One of the most important of the unanswered questions relating to both the two earlier investigations was the extent to which selective factors had been operative in placing a given child in a given home. As a preliminary to her main study, Leahy therefore undertook an analysis of all the available information on file in the Children's Bureau of the Minnesota State Board of Control on the disposition of the children of unmarried mothers born in Minnesota during the period from 1918 to 1928. Since the laws of Minnesota require separate registration of all illegitimate births, regardless of the social class of the mother, and since every care is taken to insure complete privacy of these records and to protect the unmarried mother from undesirable publicity, it is believed that not more than a very small percentage of the total number of illegitimate children born in the state escapes registration.

Records of 11,742 unmarried mothers showed a total of 9,973 children who survived to the age of two years or longer and for whom dependable information as to placement was obtainable. Of this number, 4,213 were retained by their own mothers and 2,875 were placed in adoptive homes. The remainder were cared for in various ways, chiefly through boarding homes. The two groups just mentioned constitute the chief material for the study of selective placement.

The median age at placement for the adopted children was 6.9 months; 25 percent were placed by the age of 3.25 months. Selective placement would therefore be based largely or entirely on known factors in the family background.

Information on these heads was not available for all cases, especially for the true parents. Data on education, occupation, and so forth had been recorded for 90 to 100 percent of the foster parents and for about 60 percent of the true mothers. This fact must be taken into account in estimating the possibility of selective placement. The degree of selection within the total group is probably less clear-cut than that in the group for whom more complete information about the true family background was available from the records, although it must not be forgotten that a large proportion of these children were born in rural communities and small towns where the child-placing agencies were likely to rely on their personal knowledge of the girl and her family in deciding upon the disposition of the child, so that ultimate placement would thus frequently depend upon many factors not entered in the formal record.

Absence of precise recorded data cannot be taken as evidence that the omitted facts were not known to the child-placing agencies or that these facts were not taken into account in placing the child.

Among those for whom data were given, there was no difference between the adoptive and the retained group in respect to age of mother at birth of child or racial background of true parents, but the former (adoptive group) exceeded the latter in respect to mean education and occupational level of both (true) parents. The parents of children placed for adoption at three months or younger likewise exceeded those of children placed at a later age in both educational and occupational status. The differences were most apparent at the upper extremes. Roughly speaking, Leahy's data seem to indicate that the chances that the parents had completed the high school or better and that their occupational classification was in one of the two upper groups of the University of Minnesota, Institute of Child Welfare Occupational Classification were about twice as great among the children placed for adoption as among those retained by their parents, and that the odds were again doubled if the adoption had taken place by the age of three months. This is a fact of much importance in view of the repeated finding that adopted children average somewhat above the norm for the general population in tested intelligence and the finding that those adopted at an early age commonly excel those placed at later ages.[1]

Coefficients of contingency between the educational and occupational levels of each of the true parents (and even the true paternal and maternal grandparents of the children) and the same measures for the adoptive parents are all positive, and their values are typically about four times their standard errors. Comparison with such data as are available for the general population suggests that the younger group of adopted children are at least not inferior to the average in mental potentialities, as far as these can be judged from the family history data reported. They may even be somewhat above the average. Selective placement on the basis of family history is shown to exist.

---

[1] See summary by Speer in Part II of this Yearbook.

After the completion of the preliminary study, a controlled investigation (13) of the later status of a selected group of 194 adopted children was carried out. All children were of illegitimate birth, all had been placed in adoptive homes before the age of six months, and all had been legally adopted. At the time of the investigation all were between the ages of 5 and 14 years, a range selected because it is the period for which the 1916 Revision of the Stanford-Binet is best adapted. Only children reared in communities of 1,000 or larger were used, and only those cases in which both adopted children and adoptive parents were white, non-Jewish, and of north-European extraction. Each of the adopted children was matched with an own child with respect to sex, age, occupation, and education of (foster) parents, who met the same restrictions with respect to racial background and residence as were laid down for the experimental group.

For the details of testing procedure, the reader should consult the monograph cited. Here it is sufficient to say that the conditions of matching were rigidly enforced, but it was found that the homes of the adopted children ranked slightly higher than those of the control children on the Leahy Scale for Measuring Urban Home Environment.[1]

As for the results, the difference between the two means is slightly less than one-third of a standard deviation of the distribution of control-group scores. Leahy suggests that the difference may be at least in part attributable to the smaller size of the families of the adoptive parents (in the majority of cases the adopted child was the only child), which would leave more funds for the acquisition of material possessions and more time for participation in community activities, both of which weigh rather heavily in determining the score.

The two groups of children earned almost identical means on the 1916 Revision of the Stanford-Binet, but the distribution of IQ's for the adopted children is sharply curtailed at the lower extreme and to a lesser extent at the upper extreme. There are nine control children and only one adopted child with IQ's below 85; there are eleven controls and five adopteds with IQ's of 135 or higher. The rarity of very low IQ's among the adopted children may be either the result of very careful selection of adoptive homes on the part of child-placing agencies, a practice known to exist, or of the state law prohibiting the placing of children of feebleminded parents for adoption until an age has been reached at which the child's intellectual level can be determined. Those

[1] Scores on this scale were not used as a criterion for matching, since standardization of the scale had not been completed at that time.

who stress the importance of environmental factors as determinants of child intelligence will undoubtedly lean toward the first explanation; the hereditarians will prefer the second. The mean IQ of both groups is above the average of the population as a whole (for the adopted children, it is 110.5, S.D. 12.8; for the controls, 109.7, S.D. 15.4).

Correlations between the IQ's of the control children and such factors as parents' standing on the Otis Intelligence Test, parents' education, and the various sections of the Leahy Home Rating Scale are in the neighborhood of +.50. For the adopted children, the correlations (corrected for inequality of IQ variability) average in the neighborhood of +.23. The difference between these correlations is illustrated further by the fact that for the control group, where the correlation between occupational class of parent and IQ of child is .45, the mean IQ's decrease steadily and markedly as you go down the scale of occupational classes. For the adopted children the correlation between these measures is only +.14, and the regressive differences are correspondingly slight.

Further indication of the non-determining effect of the home environment upon intellectual variance is to be seen in the low correlations between intelligence quotients of unrelated children reared in the same household, all of which closely approach zero.

It is unfortunate that in her conclusions Leahy reproduces the questionable pattern previously set by Burks. She attempts to make a precise quantitative estimate of the proportion of the intellectual variance of the foster-children that can be accounted for on the basis of environmental factors, overlooking the fact that in so doing she is generalizing from a selected, and presumably rather small, part of a non-homogeneous total. She might perhaps be warranted in saying something like this: "The variances in home environment indicated by different levels of education, occupational status, or scores on the Otis Intelligence tests of the foster parents do not, when considered separately, account for more than 4 percent of the variation in intelligence quotients found to exist in a group of foster-children placed in these homes before the age of six months."

Here it should be noted that Leahy made no attempt to ascertain the *combined* influence of the factors studied, either by multiple correlation or by the newer technique of the discriminant function developed by Fisher.[1] However, this is not the main point at issue. The real question

[1] The reason for this omission is doubtless the fact that the former method is laborious and in view of the small size of the individual correlations with child

is, *How completely do the particular measures used sample the entire range of pertinent factors in the home environment?* One is not warranted in generalizing from a sample to a whole unless the sample is a truly representative measure of the whole. One is not justified, therefore, in making final pronouncements about the relative contribution of heredity and environment *in toto* on the basis of the measurement of a very small number of factors in the environment. The most significant features of the environment may not have been measured at all.

Another unmeasured source of variance is to be found on the side of inheritance. Child-parent resemblance tells only a part of the story; there is the more remote ancestry that must also be considered. This principle, which is well recognized by geneticists, has been pointed out by Burks (4).

There is also the variance due to chance, by which we mean whatever combination of momentary conditions of attention, health, motivation, and what not causes a child's IQ to vary from one examination to another. Since no mental test is wholly free from the influence of variable factors of this kind, there is obviously an appreciable portion of the total variance that need not be accounted for at all, since it is not 'real,' at least in the sense that it does not constitute a stable portion of the measured variability of individual intelligence.

While I question the accuracy of her final conclusion, since it seems to me that techniques thus far developed are not sufficiently exact nor the possibilities of the occurrence of conditions that may affect the two groups in an unequal degree sufficiently controllable experimentally to justify more than a pretty general statement, it must nevertheless be admitted that Leahy's data offer little positive support for the assumption that the intelligence of the children had been materially changed by transference to a foster home. Such a claim would involve the unlikely assumption that the homes of those adoptive parents who ranked lowest in the socio-economic scale were essentially comparable in their effect upon child intelligence with homes in the highest professional classes, although among the paired controls the difference between these extremes was great enough to produce a mean difference in the intelligence of the children of 16.5 IQ points. Compared to this, the adopted

---

IQ and the high intercorrelations of the environmental variables, it seemed unlikely that the findings would have been significantly modified had the multiple correlation method been tried. The Fisher technique had not been developed at the time the study was made. Had either of these methods been used, it is likely that the estimation of 4 percent would have been slightly increased.

children at the extremes of the occupational distribution show a difference of only 4.8 IQ points, some of which can almost certainly be accounted for by differential placement. As a matter of fact, Leahy shows that such factors as education and occupation of the true mothers correlate about as highly with the various measures of the foster home and of the foster parents as do the IQ's of the foster children.

2. The Study of Orphanage Children by Lawrence

Even before the appearance of the *Twenty-Seventh Yearbook,* an important investigation on the later development of foster children removed from their true parents in infancy and early childhood had been undertaken by Lawrence (11) in England, though the results were not published until 1931.

This study differs from those previously described in that it is based upon orphanage children whose environment, at least in its gross aspects, is similar for all cases although the inheritance factor varies considerably from child to child.

In this study, Lawrence posited, and sought to answer, the following seven questions:

1. Is the small correlation between child intelligence and socio-economic status of parents that has uniformly been found to exist among children reared in their own homes also present among children who have never lived with their parents?

The answer to this question was sought through a study of the Stanford-Binet IQ's of children in a single orphanage that admitted only first-born illegitimate children less than one year of age (average age of admittance 6 months). Further conditions of entrance were that the mother should in all cases give the name and occupation of the child's father, that no case should be taken unless the father had deserted, and that once the child was admitted to the orphanage the mother should renounce all further claim to it. Until about the age of 6 years, the children were " boarded out in approved cottage homes in the country." The cottagers who received them were usually " agricultural laborers of the better type." At about the age of 6, all children were returned to the headquarters of the organization where they remained until about the age of 16, attending their own school within the orphanage and leading a secluded life within the walled grounds. None, apparently, was placed for adoption. The mean age at the time of taking the Binet test was slightly over 10½ years; the mean IQ for 231 boys was 101; that for 153 girls, 99.

The occupations of the true fathers were known for 160 of the boys and 109 of the girls. These occupations were graded according to a five-fold classification and the correlation ratios between child's IQ and occupational class were computed. These were found to be + .22 ± .05 for the boys and + .26 ± .05 for the girls. The mean IQ's of the children in the top and bottom classes respectively were for the boys, 102 and 92; for the girls, 100 and 94.

The older children were also given the Simplex Group Test of Intelligence in order that the correlations with occupational class might be compared with those found for London elementary-school children living in their own homes. On this test the correlation ratio for 103 boys was + .26 ± .06 and that for 82 girls was + .25 ± .07. For 211 London elementary-school boys the correlation ratio was + .27 ± .04; for 238 girls, + .22 ± .04.

As a check I have recalculated these coefficients by the Tschuprow contingency method (51), which is better suited to material of this kind. This coefficient, unlike the ordinary Pearsonian contingency coefficient, varies from 0 to 1.00 regardless of the number of classes or the system of classification. The results so obtained check very closely with the correlation ratios obtained by Lawrence, as shown in the following

	Simplex Test				*Stanford-Binet*	
	London School Children		*Orphanage*		*Orphanage*	
	Boys	*Girls*	*Boys*	*Girls*	*Boys*	*Girls*
Number	211	228	103	82	160	109
Contingency coefficients	.36	.20	.20	.26	.24	.28
Correlation ratio	.27	.22	.26	.25	.22	.26

tabulation. None of the correlations is very high, but all are approximately as high for the orphanage children who had never seen their parents as for the public-school children.

2. The second question raised by Lawrence was: Is there any difference in the extent of child-parent resemblance between children who are removed from their own homes at different ages?

Data on this point were derived from another group of orphanages to which children of all ages and many social classes were admitted. The mean intelligence level of the children in these homes was similar to that in the first; that is, approximately 100. Here the range of social class was greater and the correlation ratios correspondingly higher. For children admitted before the age of three, the figures are: for 38 boys, + .34 ± .10; for 30 girls, + .297 ± .16. For those admitted after the age of three years they are for 57 boys + .599 ± .06; for 50 girls, + .35 ± .08. These figures suggest that longer residence with parents increases the relationship between paternal occupation and child intelligence, but the number of cases is too small and the standard errors are too large to warrant much confidence in the findings.

3. Lawrence's third question was phrased as follows: If children from bad homes are put into an improved environment, does their intelligence increase in proportion to the length of time they have been in better circumstances?

By way of answering this question, those children were selected who had come from homes " so bad that life in any kind of institution would be preferable." These children were divided into four groups according to the proportion of their total life span that had been spent in the institution, a method judged to be preferable to the use of absolute length of residence, since it takes account of differences in chronological age. The number of cases involved (both sexes) was 293. A very slight difference in favor of those who had spent more than 40 percent of their lives in the institution was observed (2.1 IQ points for the girls and 3.1 for the boys), but the variability was too great to warrant much confidence in this finding, especially in view of the fact that the means for all four residence classes shift about in a haphazard fashion; it is only when the cases are combined into long- and short-residence groups that any difference at all can be detected.

4. The fourth question was: When children remain in their own homes, does their likeness to their parents or to their parent's class increase with age?

It is pointed out that such an increase is susceptible of two interpretations. It might be due to progressive influence of the environment, but also it might depend upon genuinely inborn factors that ripen gradually. Thus a comparison with children who have had no contact with their own parents is needed in order to cancel the first possibility. The data from children of the London Elementary Schools were therefore compared with those from the first orphanage group. For the London children it was found that the correlations did show an increase from about .30 at age 10 to around .45 at age 13 for the girls and from .23 to .37 for the boys of corresponding ages. In the orphanage an increase with age was also found, but of smaller magnitude, from approximately .22 to .28 for the boys and .21 to .25 for the girls. However, the figures are not strictly comparable, since the age-groupings differ for the two groups and the grouping of all cases into two age levels (necessitated by the small number of cases in the orphanage group) would tend to flatten out the rise in the curve. The finding of an age trend is, however, of some interest in view of the points recently raised by Anderson (1) and the results of the study of nursery-school training by Goodenough and Maurer in Part II of this Yearbook.

5. Lawrence's fifth question was: Are children in a uniform environment more alike than those remaining in their own homes?

The answer was sought in a comparison of the coefficients of variability of the IQ's earned on the Simplex Test by the London school children with those of the children in the first orphanage. A slight difference in the expected direction was found. For the boys the figures are: 13.9 for the orphanage and 15.4 for the London school children; for the girls, 12.9 and 14.0. These differences are small in comparison with what might be expected if environment were the main determinant.

6. The answer to the sixth question, which had to do with the possibility of a difference in the findings if a non-verbal instead of a verbal test were used, appeared to vary with the test. For the orphanage girls, however, the correlation between occupational class and performance on non-verbal tests tended to be higher than that established by the use of verbal tests; for the boys it was positive but lower. No explanation for the sex difference could be offered.

7. The seventh question was concerned with the possibility of the findings having been affected by any of a number of factors, such as differences in health, nutritional status, age of parents, and the like. Results were all essentially negative.

Lawrence's study is of particular interest because of the clear and straightforward manner in which the questions to be investigated were stated and the use of orphanage children as subjects, which has the great advantage of reducing the number of variables in a problem that at best is extremely complex.

It is sufficient to quote her general conclusion:

> . . . The discovery of a correlation between the intelligence of children and the social class of their parents when they have never seen those parents is fairly conclusive evidence that the correlation so generally found for children in their own homes is not mainly due to the direct social influence of the home, but is a genuine biological fact. The association is on the whole rather smaller, however, in the case of institutional children

and there is little doubt that environmental conditions have some weight in influencing the response to tests. (p. 70)

It will be noted that Lawrence is cautious about assuming that the 'response to tests' is identical in its significance for these children who have been reared under conditions so different from those of the generality with that which has been found for other children. A test is appropriate for a given group only if it constitutes an adequate sampling of the abilities of that group. Whether or not lessened suitability of the test for measuring the ability of the children has played a part in the slight reduction of the correlations with family background for the institutional children is conjectural, but certainly it is a possibility that should not be overlooked.

3. The Study by Lithauer and Klineberg

A short article by Lithauer and Klineberg (14) on changes in IQ after a period of residence in the Hebrew Orphan Asylum is of interest only as it provides one more illustration of the hazards of accepting 'gains' at their face value. The study need not be described in detail, since the procedure, among other questionable practices, included the mixing together of children in boarding homes and those in the institution and of tests given at different ages and after various intervals of time. There was no control group. The results, taken at their face value, indicate a mean increase in IQ of approximately 6 points. The amount of the increase was unrelated to length of residence; a period of three months apparently accomplished as much for a child as one of several years. However, the correlation between amount of gain and age at initial test was —.42 for the 120 cases studied.

Now, anyone understanding the principles involved and acquainted with the circumstances of testing could have predicted the character of these results in advance. First, not all children in the institution were retested. The authors point out that the reasons for retesting were such as to select the children with initially low scores. The mean initial IQ of the total population was 94, of the retested group 82. The mean *gain* of 6 points is thus about what would be expected on the basis of regression, inasmuch as the correlation between test and retest was only .76, and the greater gain of the younger children can be explained on the same principle, for the retest correlation has uniformly been found to be lower for younger than for older children; hence the regressive 'gain' for that group would be larger in amount.

4. The Iowa Studies of Foster Children

The studies of foster children that have attracted the greatest amount of attention during recent years are those from the University of Iowa carried out by Skeels and Skodak (24, 25, 26, 29, 30). The net results of these studies are so startling in their implications as to merit the closest attention. On the surface, they seem to indicate that biological factors have no bearing whatever upon the later mental development of children who are removed from their parents at a sufficiently early age. This, at least, appears to be the earlier point of view of the authors, as far as one is able to judge from their published articles. For example, Skodak, in her monograph (30), says:

> . . . if there is an hereditary constitutional factor which sets the limits of mental development, these limits are extremely broad. Within these, environmental factors can operate to produce changes which for ordinary purposes may represent a shift from one extreme to another of the present distribution of intelligence among children. (pp. 132–133)

The evidence presented in support of this and similar statements both in scientific and popular journals, as well as in radio addresses and newspaper articles,[1] is less convincing than the extreme nature of the claims would appear to demand. They are for the most part based upon a study of the later development of a group of 154 children placed in adoptive homes before the age of 6 months and of an additional group of 65 children placed between the ages of 2 and 5½ years, together with the large number of cases more recently added to the group.[2] In addition, the orphanage project mentioned in an earlier section of this chapter is cited as corroborative evidence.

Inasmuch as the monograph by Skodak (30) provides the most complete data on the development of children from foster homes thus far reported from the Iowa laboratory, we shall confine our discussion to this, disregarding the other publications, which include fewer cases and in which the data are less completely analyzed.

Skodak's monograph has the merit of being clearly written, and the statistical treatment, while not always of a kind to be commended, is at least straightforward. Nevertheless, there are strange contradictions between (*a*) her recommendations with respect to child-placement policies, (*b*) her interpretations of the statistical findings of the

[1] For a summary of some of the most radical statements, the reader is referred to the article by McNemar (15) previously cited.

[2] See Skeels, Part II, Chapter XX, of this Yearbook.

study, (*c*) her summaries of these findings (which for the most part are accurate, though not always complete, because facts that are important for interpretation are sometimes overlooked), and (*d*) the complete findings reported in the body of the monograph. Following are examples:

> Since the relationship between true-family background and the child's mental development is approximately zero and since knowledge of the child's true-family history has no predictive value, the use of true-family histories as a basis for the placement of the child has little or no justification. (p. 133)

> The coefficients of correlation between the child's first IQ and true-parent education ranged from .04 to .11 and on second examination from .28 to .33. While the coefficients are less than frequently reported for parents and children in their own homes, they are slightly higher than similar correlations for the foster parents. (p. 104)

On pages 67 and 78 of Skodak's monograph the *r*'s between child IQ (second test) and education of father, mother, and mid-parent are reported. These coefficients are not corrected for range. The S.D.'s

	True	*Foster*
Father	.29 ± .07	.16 ± .05
Mother	.28 ± .05	.19 ± .05
Mid-parent	.35 ± .07	.18 ± .05

of the reported education of the two groups (pp. 40, 48) are as follows (mid-parent S.D. not stated):

	True	*Foster*
Father	2.5	3.8
Mother	2.2	2.722

When the correlations for the true parents are transmuted into their most probable values for variability corresponding to that of the foster parents,[1] the figures obtained are .38 for father and .33 for mother.

Skodak's Table 10 (p. 77) gives the following relationship between true mid-parent education and child's second IQ:

[1] See T. L. Kelley. *Statistical Method.* (Macmillan: New York, 1923) pp. 323–325.

	Eighth Grade or Less	*High School*	*Beyond High School*
Number	21	56	10
Mean IQ	104.2	111.0	120.5
S.D.	12.5	11.2	14.2

Corresponding data for the foster parents are not given.

Warrant for the following is hard to find:

> . . . Continued residence in an inadequate home results in a decline in mental level with increase in age, whereas placement in a superior home, even during the later preschool years, results in an increase in mental level compared to expectation from the information on the true parents. (p. 132)

Considering first the 154 children placed at 6 months or earlier, Skodak's Table 13 (30, p. 98) shows that when a comparison is made between (*a*) children who *gained* 6 points or more between first and second examination, (*b*) those who *lost* 6 to 15 points, and (*c*) those who *lost* 16 points or more,[1] the mean occupational levels of the true fathers of the three groups, as rated by the Minnesota Occupational Classification,[2] run as follows: Group A, 4.9; Group B, 5.6; Group C, 5.8. For the foster fathers, the corresponding figures are: Group A, 3.3; Group B, 3.4; Group C, 3.2. For mid-parent education, in terms of school grades completed (true parents), the figures are, in order as

[1] The 78 children whose shift in IQ did not exceed 5 points are not included in Skodak's table.

[2] Averaging unequally spaced groups, such as those of the Minnesota Classification, is, of course, mathematically indefensible. The classificatory scheme was devised as a rough means for comparing samples, for which purpose it has proved to be useful. It is not a measuring instrument and the numbers attached to the various classes have no mathematical significance apart from the fact that they do form something of an intellectual hierarchy. This is shown by the fact that both for the men engaged in the various occupations and for their children, a steady decrease in mean intelligence-test score is found as one goes down the list of classes, with the single exception of Class IV (Farmers). The original classification included six urban divisions only. When it was decided to add the rural group, a question arose as to where it should be placed. Because of the wide range of ability represented by the farming group, it was tentatively decided to place it at the mid-point of the classification. Since the various classes are categories only, its precise location is not a matter of serious consequence when only interclass comparisons are to be made. However, it now appears that in order to fit into the hierarchy, the positions of Classes IV and V should be interchanged.

above, 11.6, 8.6, and 9.8; for foster parents, 12.0, 11.3, and 11.5. On a home-inventory scale developed for the purposes of the study, the figures (based on the foster homes) are 85.2, 85.0, and 81.7. The standard deviation of the inventory score for the entire group is 14.3; for the three subgroups, 10.3, 10.0, and 16.4. One would be quite ready to agree that none of these figures has much meaning, since they are based upon changes between testings the first of which was given at so early an age as to be non-predictive of later mental status in any case. Nevertheless, it should be clear that these data at least offer little support for the statement quoted.

A further line of supposed evidence offered by Skodak is readily perceived to be an instance of statistical misinterpretation. On pages 99 to 100 she compares the changes in IQ from first to second examination of those who at the time of the *second* examination tested 115 or higher with those who then tested 104 or lower. The usual regressive shift is seen when the mean scores of the two groups at the time of the first examination are compared with those of the second. The group finally testing 115 or above had an initial mean 1.8 points lower than that on the second test; the one group with IQ's of 104 or below originally averaged 10.3 points higher. The fact that in this case the regression works in the reverse of the usual direction because the selection was made on the basis of the second test rather than of the first should not mislead anyone conversant with the general principle.

Differences between the foregoing groups in respect to education and occupation of true parents, and the " foster-home-inventory score " are all fairly clear-cut. Taken at their face-value, the differences between these background factors for the foster parents seem to be slightly, though probably not significantly, greater than those for the true parents. The problem is complicated by selective placement (Skodak states that the r between mid-parent education of true and foster parents was $+.30$) and more particularly by the fact, previously pointed out, that the calculation of means from the irregularly spaced occupational categories (and probably for reported parental education [1] as well) is of questionable validity.

---

[1] While no quantitative formulation of this matter has been made, as far as I am aware, everyday observation would lead to the opinion that the difference, for example, between completing two years of high school and completing three is in all probability less than that between completing three years of high school and remaining to graduate or between completing high school and taking one year of college work.

The most striking relationship between child IQ and foster-home environment is seen when the results of the second test are compared with the scores earned on the home-inventory scale. The mean scores on this scale decrease regularly with each 10-point decrease in IQ, from 104.5 for the children with IQ's of 140 or higher (3 cases) to 67.0 for those with IQ's below 90 (4 cases). The correlation between the inventory score and IQ for the entire group was .49 ± .06. While this is a highly important finding, one should be wary of accepting it at its full value. Examination of the items shows that most of them are highly subjective in character, which renders the scores easily susceptible to the effects of unconscious bias. That such an influence may easily have existed is evident from the description of the conditions under which the inventory was taken. We are told that

> The scale was administered to 103 children in the 0- to 6-month placement group, who were between the ages of three and six at the time the scale and the last examination were administered. The scale was given by the writer in all except four cases . . . in every case the examination preceded the interview [1] with the parents. In discussing the results of the examination the necessary questions, with some explanation of the purpose of the scale in learning the usual practices of homes, were woven into the conversation. (p. 71)

Obviously no one can be certain that the relationship between child IQ and home rating was affected by the manner of securing the information, but it is evident from the quotation just made that the conditions strongly favored such an effect, since the inventory-taker knew the test results and rated the home while discussing these results with the parent.

The second part of the Skodak monograph, which is concerned with the development of 65 children placed in foster homes between the ages of 2 and 5½ years, will not be reviewed in detail since in the main we should be repeating points already made. The findings from this part of the study are further complicated by the fact that the children had spent varying lengths of time in their own homes and in different institutions before the study was begun. All had at some time been residents of the Iowa Soldiers' Orphans' Home, the institution described in the study of orphanage children by Skeels, Updegraff, Wellman, and Williams (28). On preplacement tests, the mean IQ for the entire

[1] That is, the securing of the informational data of which the scale is made up. Most of the items in the scale are based upon parents' statements.

group was 98.5.[1] A year after placement, the mean IQ for the entire group of 65 had risen to 104.2 — a change not significantly out of line with that reported in Part II of this Yearbook for both experimental and control groups in a number of studies on the effect of nursery-school training. A correlation of $-.36 \pm .07$ between IQ at placement and gain after placement "confirms the tendency for the initially lower IQ's to rise on subsequent examination and for the higher IQ's to remain constant or to lose slightly," as Skodak says; but it also confirms the existence of the statistical fallacy in such computations pointed out by Thorndike (39) and by Thomson (38).

The brief review by Skeels (Part II, Chapter XX, of this Yearbook) on the changes in IQ of originally feebleminded children placed in an institution for the feebleminded and cared for by moron girls raises a number of interesting questions. The fact that first strikes the reader's eye is the extremely large change in mean IQ reported for both the 'experimental' and the 'contrast' groups. The gain of 27.5 IQ points by the former is so far in excess of the 4.6 points brought about by the University of Iowa experts in the Davenport experiment or the 5.5 IQ points accomplished through placement of children old enough for preplacement tests in adoptive homes of superior caliber [2] that one wonders whether it might not be a good idea for the rest of us to take lessons in child-training from the morons. Certainly they appear to have done a better job than any one else has been able to accomplish!

The extremely large mean loss in IQ by the 'contrast' group retained in the orphanage is also puzzling. This group is said to have dropped from the initial mean of 86.7 to a final mean of 60.5; yet in the experiment by Skeels, Updegraff, Wellman, and Williams the mean loss in IQ for the control group was only 4.6 points as compared with this loss of 26.2 points for Skeels' 'contrast' group. Inasmuch as both of these groups were residents of the Davenport orphanage throughout the period of study, it appears that some unreported factor of selection must have been responsible for the difference. It should be noted, however, that at the time of the first tests upon which the diagnoses of mental deficiency (leading to placement in the institution for the feebleminded) or of approximate normality (leading to temporary retention in the orphanage) were made, the mean age was only 19.4 months.

[1] This is on the basis of the last test given before placement. Twenty of the children had been given a test previous to this. The mean IQ earned on this test was 97.0. The small difference can best be attributed to chance.

[2] See Skodak, Ref. 30, Chapter IV.

The strongest argument for the effect of placement in foster homes upon the intelligence of children is to be found in the high mean IQ of these children. In a later report Skeels [1] states that of 306 children committed under 6 months of age and placed for adoption before or after the age of 6 months, the mean IQ was 117.2, with standard deviation of 14.5. The mean age at time of testing was 22.7 months. Inasmuch as Skodak shows that, within her group of 154 cases, the mean IQ dropped steadily from 120.4 for tests given at one year [2] to 108.2 for those given at 6 years, the later development of these children is a question of much interest.

It should be noted that the high mean IQ of the foster children studied by Skodak and Skeels is not a unique finding. Working with older children, Burks obtained a mean IQ of 107.4; Leahy of 110.5. Skodak's figures show a mean of 110.6 at age 5 and 108.2 at age 6. Evidently, except for the difference in the age of the children, the Iowa figures are not seriously out of line with those reported by others; the main difference is in the interpretation. The Iowa investigators consider that this mean is far above genetic expectation; Burks and Leahy regard the increase resulting from the improved environment to be positive but small. Burks' estimate is from 6 to 8 points of IQ for an environmental difference of one standard deviation; Leahy's estimate of the total contribution of environment is a shift of 3 to 5 IQ points.

Obviously, for lack of adequate data these estimates cannot at present be subjected to crucial tests. We can, however, apply certain checks that may help to throw light on the problem.

In their report of the standardization of the 1937 Revision of the Stanford-Binet, Terman and Merrill (37) give the mean IQ obtained on the new scale by children in the different occupational groupings of the Minnesota classification used by Skodak and Skeels and by Leahy. They also present a comparison of the IQ's earned on the 1916 and the 1937 Revisions of the Stanford by ages separately for children who were given both scales within an interval of 5 months or less. Of these, 158 were between the ages of 5 and 14 years, which is the age-range used by Leahy. Their mean IQ on the 1916 scale was 101.2; on the 1937 Revision 101.0. It would appear, therefore, that no great error is involved in utilizing the data from the 1937 test as given in Table 12 (p. 48) of *Measuring Intelligence* (37) to make a first estimate of the IQ to be expected in Leahy's group on the basis of the reported occupations of the true fathers.

[1] See Reference 26 citing a contribution in Part II of this Yearbook.

[2] Age taken to nearest birthday.

The mean IQ of the 89 children for whom the true fathers' occupations were known was 110.9. A weighted average, found by multiplying the mean IQ reported for each occupational level in the Terman and Merrill table by the number of true fathers at that level and dividing the sum by the total number of cases, gives 101.8 as a first approximation to the IQ expected on the basis of paternal occupation alone. However, some correction of this figure is needed. Leahy does not state the ages of the fathers in this group, but in the earlier survey she reports the mean age of the true fathers of children placed for adoption as 26.5 years, S.D. 7.5 (994 cases). The Minnesota classification is based on males between the ages of 21 and 45 years; Leahy's group presumably averaged about 7 years younger than the mid-point of the standardization group. Some allowance should, therefore, be made for the youth of these fathers, many of whom would not yet have attained their final occupational status. Just how great this allowance should be is uncertain, but comparison of the distribution of occupations in the next lower age-group given in the Census report makes it evident that the proportion of men in the higher occupational brackets is very much smaller than it is in the group from which the Minnesota percentages were derived.

Possibly a further correction is needed because of the fact that children of feebleminded parents were not placed for adoption, but the number of such cases in the general population is not great enough to have more than a small effect upon the values cited in the Terman and Merrill table. In so far as this admittedly very inadequate analysis goes, however, it indicates that Leahy's estimate of 3 to 5 points of IQ probably understates the contribution of the environmental factor. The truth probably lies somewhere between this estimate and the 9-point difference suggested by the distribution of paternal occupations.

An attempt was made to carry out a similar computation for the Iowa children, but it was found impossible to do this because of the irregular combination of Kuhlmann-Binet and Stanford-Binet tests and the absence of data for transmuting either into equivalent units of the 1937 Stanford. Had the occupational distribution by ages separately been presented, a comparison could have been made, using the Goodenough results on the Kuhlmann-Binet for the younger ages and the Terman and Merrill table for the older cases. It may be noted in this connection that the graph presented by Skodak (30, p. 65), in which the mean IQ's of the Iowa children grouped by occupation of foster parent is compared with the findings for each of these studies, is invalid, inasmuch as no account has been taken of the differences in the tests used.

Summarizing, we may say that there is evidence in the Iowa studies for some improvement in tested intelligence of children placed in foster homes at an early age over that which would be expected on the basis of home background. That the transforming effect of the foster-home placement is as great as the enthusiasm of the authors of these studies might lead the uncritical reader to believe seems highly questionable when the many sources of error in their published reports are taken into account (15). It should be noted that the children studied are for the most part still too young for intelligence tests to have a high degree of predictive value for ultimate mental status. Moreover, the fact that the IQ's show such a steady decrease with age raises a very real question as to the actual significance of these early test scores. Those familiar with the Kuhlmann-Binet Scale, the one used mostly with children under age 3½ years, will recall that the tests for the early levels consist chiefly of simple reflexes and perceptual-motor skills that have never been found to correlate highly with tests given at a later age. In the light of recent reports by Bayley, Furfey and Muehlenbein, and others, one should hesitate long before placing too great confidence in these early tests as evidences of what we later call 'intelligence.'

The statement by Skodak in her monograph (p. 104) that "the mental development of children of feebleminded mothers and the most inferior true-family backgrounds is indistinguishable from that of children whose mothers are not feebleminded" is startling in its implications. If true, a practicable vocation for the feebleminded girl has at long last been found. For if there is truly no difference in the mental caliber of the infants produced by her and those produced by women of high intelligence, it may be argued that it is socially and economically wasteful for normal women to devote their time and strength to the bearing of children. Perhaps we should turn our institutions for feebleminded women into breeding centers! But before doing so, perhaps we should pause to note that the Iowa conclusion is based upon only 16 illegitimate children, which means that the paternal side of the ancestry is at best questionable (the occupations of only 10 of the true fathers are reported), while the fact that the mean educational level of the mothers was 7.6 grades completed raises some question as to the correctness of the diagnosis of 'mental deficiency.'

5. Atkinson's Study

Let us now turn to the evidence from other sources that deals specifically with the question of the later development of children of feebleminded parentage reared in foster homes. The earliest of these based upon tests in present-day use that I have been able to locate is one by Atkinson (2).[1] Because of the youth of the infants at the time

---

[1] Two or three studies have appeared in which the diagnosis of mental deficiency in the parents was based either upon popular opinion or upon the Goddard

of testing, together with the fact that these infants, along with their mothers, were residents of the Women's Reformatory at Clinton, New Jersey, and thus were in close daily association with their mothers, even though both were under the direct supervision of the institution, a brief note regarding the findings of this study will suffice. The subjects were 25 white infants, ranging in age from 4 to 27 months at the time of testing. Of the mothers, 18 were said to be feebleminded or of borderline intelligence; 7 were of average or inferior intelligence. The tests used for the infants were the Gesell Developmental Schedules and the Kuhlmann Revision of the Binet. All the babies were physically normal.

Of the 13 children of feebleminded mothers, 5 were classed as " feebleminded " on the basis of the combined results of the two tests; 6 as of " borderline intelligence," and two as of " inferior intelligence." Of the 5 infants whose mothers were " borderline," 3 were " feebleminded," 1 " inferior," and 1 " average." The one mother classed as " inferior " also had an " inferior " child. The 6 " average " mothers had 1 " feebleminded," 1 " borderline," and 3 " average " children, with the remaining child unclassified as to intelligence.

Although these findings are about what a geneticist might expect, the data are too scanty and the possible sources of error are too many to inspire much confidence.

6. The Study by Snygg

The findings of a recent study by Snygg (31) are in apparent agreement with those from Iowa in that little relationship was found between the intelligence quotients of true mothers and those of their children who had been reared in boarding homes.

Owing to the fact that information on many basic points is lacking, this study is hard to interpret. It is said, for example, that " the 312 children who are the subjects of this study differ from the 1,350 others who have been examined by the agency's mental hygiene clinic only in that the Stanford-Binet intelligence scores of their mothers are known." But this may be a highly important difference. Why and under what circumstances were these mothers selected for testing? Were the children legitimate or illegitimate? If illegitimate, and the mothers were tested at or near the time of confinement, the IQ's so obtained have little significance. The information on the age of the chil-

Revision of the Binet. Inasmuch as the standardization of the latter was such that some 30 percent or more of individuals who were past the age of childhood would by its use appear to be feebleminded, it has not seemed defensible to include these studies in the present report.

dren at the time of testing is both scanty and confusing. The graphs showing the percentage of children with IQ's over 90 and below 70 grouped according to the intelligence of their mothers are said to be based upon the last test given to each child, but inasmuch as the 312 children seem to have been given 899 tests, of which only 149 were given at 3 years or later, it is evident that more than half the cases included in these graphs were tested at an age when little confidence can be placed in tests as predictive instruments. In spite of all this, it appears from Snygg's data that in the case of girls who tested below 70 (98 cases) the chance was about one in eight that the children would also test below 70, while among those who tested above 90 (75 cases) the chances of producing a child with IQ below 70 were only about one in 33.

The low correlations between mother's IQ and child's IQ cannot be compared directly with those obtained from an unselected group, as it is evident that the range of maternal IQ's is greatly restricted. If corrected for range, I doubt that these *r*'s would be much, if any, lower than would be found for infants reared in their own homes when tested at so tender an age, particularly if, as I suspect was true in this case, the mothers also were tested under conditions unfavorable to securing a true estimate of their ability. Incidentally, Snygg states that the regression was not rectilinear, yet he uses Pearsonian *r* as his measure of correlation. If he is right, then the true correlations as indicated by the correlation ratios would be higher than the *r*'s reported by him. Although this study has frequently been quoted as providing support for the Iowa findings, it appears like a slender reed to offer as a prop for conclusions so weighty.

7. The Study by Wells and Arthur

Another investigation on the later development of the children of feebleminded parents placed in foster homes has been reported by Wells and Arthur (49). The definition of ' foster home ' used in this study is not entirely clear; the authors merely state that " none of the foster-child group had been placed for purposes of adoption." Since they add that " all the foster homes were licensed by the state and supervised by some one of the child-placing agencies participating in the study," it seems probable that boarding homes, rather than adoptive homes, are meant. The subjects of the investigation were 100 children, one or both of whose parents were feebleminded,[1] who were placed in foster homes at an average chronological age of 5 years, 6 months, and a control group of 100 children, also of feebleminded parentage, who were left

[1] The criterion of feeblemindedness in the parents was simply that of a Binet test record (form not stated) of 75 or below.

in their own homes. The mean chronological age of the second group at the time of the first testing was 6 years, 7 months. The foster-home children were first tested either "during residence with their natural parents or within two months after removal from them and were retested after a sojourn of at least two years in a foster home." The mean chronological age at retest was 10 years, 5 months. It is stated that

> . . . in only five cases had foster-home placement been the direct result of the child's own behavior. The reasons given for the removal of the other 95 cases from their own homes included institutionalization of one or both parents, delinquency of parents, death of one or both parents, desertion, drunkenness, or neglect. (p. 280)

We are justified, therefore, in questioning the adequacy of the initial tests given to this group, for it is evident that emotional factors likely to lower the initial test results below their true value may have been present with these children to a much greater extent than would be true of the control group, for whom such conditions were at least not sufficiently extreme to necessitate removal of the children from their homes. Moreover, the fact that an unspecified number were tested "within two months after removal from their own homes" increases the likelihood that the initial scores might not adequately represent their true ability.

The children of the control group were retested at least 2 years after the initial testing. The mean age of this group at the time of the retest was 12.

The tests used were "the Kuhlmann, the Kuhlmann-Binet, or the Stanford-Binet,[1] except one retest with the Kuhlmann-Anderson given as an individual test." Results of the testings are given in terms of IQ and the Heinis Personal-Constant, or PC, which, since it was empirically derived from retest data by the Kuhlmann-Binet and was designed to correct a tendency of the scores of inferior subjects to decrease with advancing age, is probably the preferable measure to employ with that test. The mean personal-constant of the foster-child group was 85 on the first test and 90 on the second; a gain of 5 points. Whether or not this represents a statistically significant difference cannot be determined, for the correlation between first test and retest is not given.

[1] The distinction between the Kuhlmann and the Kuhlmann-Binet is not made clear. Possibly the former is the new Kuhlmann Individual Test, which, according to its author, is of the Binet type but employs none of the items of the original Binet. I have seen no warrant for combining scores from this test with those of the earlier Kuhlmann-Binet or for treating the latter as if it were identical with the Stanford.

The mean PC on the initial test of the control group was 87.6; on the retest, 87.4, or a loss of 0.2 points. As far as mean rating is concerned, it is evident that the control group showed practically no change in status; the foster-child group, a slight tendency to increase. This may have been either the result of improved environment or of improved coöperation resulting from lessened emotional tension. No data on the range or frequency of individual changes in either IQ or PC are given, but the mean gain or loss for groups of varying levels according to the initial test is given. These tables show the usual statistical regression — cases originally testing high appearing to lose, those originally low to gain — and are of no consequence for our discussion.

In addition to inequality of initial testing conditions, the groups were dissimilar in a number of respects that might be expected to influence the results. Tests given to young children are known to have far less predictive value than those given to older children. Only 28 of the 100 children in the control group were tested before the age of 5. The number of the foster children who were tested before 5 is not stated, but since 38 of them were placed before that age, apparently the number tested before that age was not less than this figure.

The fact that the younger group of foster children showed a mean gain in PC of 11.00 points, while those placed after the age of 5 gained only 0.90 PC points, is hard to interpret without knowing whether or not the initial standing of the two groups was truly similar. If the younger children were more upset than the older ones by the change in environment, such a result is quite in line with expectation. There is also the unanswered question of varying intervals between test and retest.

Since these were all children of feebleminded parents, the distribution of test scores is of interest. This is given only for the initial test and in terms of a four-way classification. More precise figures would have been desirable. Because later tests give a better approximation to the maturity level than those given at younger ages, results based on the second test would have been of greater interest. Taking the results as they stand, the number of cases at the two extremes is the same when considered on the basis of IQ as it is when the PC is used as the interpretative measure. In the control group there were 30 cases initially testing below 75 and 17 at 95 or above. In the foster-child group there were 35 below 75 and 14 at 95 or higher.

In their summary, the authors raise a number of questions, to many of which tentative answers can be given. First, they are concerned to know why, when 50 of the control-group cases lost on retest, the other 50 did not also lose. Also they ask why 38 of the foster children failed to gain. Unreliability of the measuring instrument is certainly the most likely answer to both these questions.

The regressive character of the gains and losses also puzzles them. One wonders how so many investigators have managed to escape an acquaintance with so basic a principle of elementary statistics.

All in all, it does not appear that the study adds much to our knowledge of the possible effect of foster-home placement upon the intelligence of children. The number of uncontrolled variables is too great. Moreover, the age of the children at time of placement was beyond that which most investigators have considered optimal for effecting a change. There is no way of knowing whether or not the small change reported for the foster children represents a genuine improvement in ability or, on the other hand, whether the change might not have been greater had the children been placed at an earlier age. Finally, while the foster homes in question were presumably superior to those from which the children came, it is highly unlikely that they were comparable to the usual type of adoptive home described in the studies by Leahy, Skodak and Skeels, and others. Just how superior these homes were to those of the children in the control group is a question that remains unanswered.

8. The Study by Speer

The study by Speer (32), reported in Part II of this Yearbook, offers the strongest support for the environmentalist point of view that I have met. On the technical side, the study does not appear to be open to serious criticism. His conclusion that "there is no reason why physically normal children of feebleminded mothers may not be placed for adoption, from their own homes, if this is done before the third birthday" may to some seem a bit hasty, in view of the fact that it is based upon only 12 cases placed at this age and tested, presumably, before the age of 5.[1]

The most perplexing thing about Speer's contribution is the statement that reads (quoting from his manuscript):

> . . . The children of both our groups [experimental and control] come from homes that are definitely of low economic status. The only meas-

[1] The latter point cannot be ascertained with certainty because the data are so presented that it is impossible to relate the table showing age at placement to the one showing age at testing.

> urable difference in the homes of the two groups is the mental status of the mothers. . . . The differences between the median IQ's of the two groups might be taken as a measure of the effect of the presence of the mentally deficient mother. It will be noted that the difference becomes greater, the longer the children have been in their own homes.

Elsewhere it is stated that "the histories reveal very few differences between the fathers of the two groups. In both groups, the fathers are described as irresponsible, alcoholic, epileptic, mentally ill, or venereally diseased."

To me this seems like placing an excessive weight upon the part played by the mother in determining the character of the home environment. Unless mental deficiency is contagious, like measles, one would expect that the other members of the household, the school, and other social influences would play a part in determining the outcome. These things are said to be equal for the two groups; yet the children whose mothers were feebleminded became progressively more and more backward with advancing age if they remained in their own homes, whereas those whose mothers were normal showed little or no such tendency in spite of the other adverse home conditions. One might think that if environment were the main determining factor, the influence of the feebleminded mother might be felt most heavily during early childhood when the child is most dependent upon her for care, and that the part played by the mother alone would be of less significance later on, say after school age. This, however, does not seem to be the case, if Speer's data can be depended on. Evidently the point is one that calls for further study.

9. The Study by Reymert and Hinton

In this connection it may be well to call attention to a similar trend in the article by Reymert and Hinton (19). In their study the IQ's of children entering the Mooseheart institution at different ages decrease steadily with age at entrance from a mean on the fifth test of 103.25 for those entering at 6 years and earlier to 97.0 for those entering between the ages of 7 and 9 years, and 91.65 for those entering after the age of 10. For the youngest of these groups there is an increase of a little over 5 IQ points during the first year after entrance; for the others, there is no consistent change in rating during a residence period of 4 years. The authors offer no explanation of the consistent decrease in mental-test standing with age at entrance. Whether it is a reflection of some unnamed selective factor or a true indication of the deleterious effect of longer stay in their own homes cannot be said from the data at hand. It would be well to have more information regarding the conditions under which children are admitted to Mooseheart.

10. The Study by Stippich

The brief study by Stippich (34), like that by Speer, is based upon a comparison of the mental development of the children of normal and of feebleminded mothers reared in boarding homes. Stippich's comparison is somewhat more direct, however, since all children had been placed in boarding homes before the age of one year. According to this study, about one-fourth of the children of feebleminded mothers reared in boarding homes since infancy earned IQ's below 75 on tests given between the ages of 3 and 12 years, and fewer than half of them tested above 90. None of the children in Stippich's control group whose mothers were of normal intelligence tested as low as 75; more than 80 percent earned IQ's of 90 or higher. Although the number of cases is not large, the differences are sufficiently clear-cut to meet the usual criterion of statistical significance when Fisher's procedure for the analysis of variance is applied. Since Stippich has shown that the boarding homes in which the children were reared were similar for the two groups, heredity seems to be the only reasonable explanation for the differences in their mental level.

V. General Summary and Discussion

1. Extreme Views of the Fixity of Inherited Intelligence

The publication of the *Twenty-Eighth Yearbook* of this Society focused the attention of educators, psychologists, sociologists, and biologists upon an aspect of social control that too many persons had taken for granted. Intellectual ability, or intellect, had been tacitly assumed to be in the nature of a gift from the gods. It was the task of educators to help the child to learn, not to increase his intrinsic ability to learn new things. The doctrine of formal discipline, although it occurred and reoccurred in the educational literature from time to time, was pretty thoroughly dispelled by controlled experiments early in the present century. With the rediscovery, in 1900, of Mendel's now famous work on the mechanisms of heredity, and the enthusiastic adaptation of these principles to the question of mental deficiency by Goddard that occurred little more than a decade afterward, the public mind, which has ever been ready to shed inconvenient responsibilities, became more firmly set than ever in the attitude that, as far as his mental development is concerned, the future of the child is on the knees of the gods. Then came the discovery that correlations between earlier and later administration of standard intelligence tests to children of school age are uniformly high. From this there arose the doctrine of the 'constancy of the IQ' and in its wake there followed a stream of applications and misapplications, interpretations and misinterpretations, of the experimental evidence such as has rarely if ever been paralleled in the entire history of psychology.

2. Facts about the Constancy of the IQ

The demonstrated facts are these: When children *of school age* are given one of the more modern revisions of the Binet Scale, such as the Stanford or the Kuhlmann revision, under standard conditions and by competent examiners, most of the intelligence quotients thus earned will show only small fluctuations upon retesting with the same scale after intervals varying from 1 to 6 or 7 years. Considering all the evidence, it is safe to say that 50 percent of elementary-school children will not change their standing by more than 5 points of IQ in either direction, while the remaining 50 percent will show somewhat greater variation.

A rather remarkable example of 'wishful thinking' is apparent in the almost universal tendency of school people to ignore the 50 percent last mentioned and to utilize the results of single tests in educational classification and guidance with a degree of assurance that would be warranted only if the entire range of variation were no greater than that shown in the middle 50 percent. The responsibility for this faulty handling of test data rests squarely on the educators; the psychologists responsible for the construction of the tests have meticulously pointed out the extent of variational tendency that may be expected. Terman, for example, who by many has been deemed the arch advocate of IQ constancy, has, from the beginning, laid much stress upon the fact that the term 'constancy' is only a relative one, that it refers to a degree of probability, not to absolute certainty. In *The Intelligence of School Children* (36), published in 1919, he is explicit on this point. He says in summary:

> Roughly speaking, the chances that an IQ will either increase as much as six points or decrease as much as four points are one in two: that it will either increase as much as twelve points or decrease as much as eight points, one in five; that it will either increase as much as eighteen points or decrease as much as twelve points, one in twenty; that it will either in-

crease as much as twenty-four points or decrease as much as sixteen points, one in a hundred and forty.[1] (pp. 148–150) [2]

Applying these figures to an average elementary school with a population, let us say, of 500 children, it is apparent that under the best conditions of individual testing, at least 100 children may be expected to show changes in IQ upon retest of as much as 10 points. In approximately 25 of these cases, the change will amount to as much as 15 points, and in 4 or 5 instances, as much as 20 points. If, in place of the highly trained examiners and careful checking of the scores for arithmetical errors that is presupposed in the foregoing figures, the tests are given by persons with only a moderate amount of training and skill, the frequency of such changes will be correspondingly greater. If the ordinary group test is used in place of an individual test, and particularly if the test used at the time of the later examination is not the same one as that used at the earlier examination, the frequency and magnitude of the discrepancies between first test and later test may be greater still.

The point of all this is that changes in intelligence quotients, even of considerable magnitude, are by no means the unique phenomena that many poorly informed advocates of intelligence testing have assumed them to be.

It should be noted furthermore that the tendency reported here for greater likelihood of increase than of decrease in standing has been substantiated by many other investigations. This so-called 'practice effect' can be attributed in part to greater habituation to the testing situation (lessened tension and anxiety, and increased self-confidence). Moreover, in a fair number of instances, particularly among the brighter children, a definite attempt will be made to find out the answers to test items failed during the first test, and some memory of these answers may carry over to a later testing.

All this should point the need for caution in interpreting such changes in test performance, particularly of small groups. Even when the changes seem to take a constant direction, the reason for the variation is not always as clear as it may seem. More recent investigations, together with more careful statistical analyses of the results have shown, too, that the chances of variation are not equally distributed among all levels of intelligence, but that variations in intelligence quotient are greater among bright children than among backward ones, and that this follows of necessity from the mathematical relations between

[1] These figures, it must be noted, apply only to school children and not to those of preschool age. At the time Terman wrote this, little or no information about the variational tendencies of younger children was available.

[2] Quoted by permission of Houghton Mifflin Company.

chronological age and the standard deviation of mental ages essential for the valid use of the intelligence quotient as an interpretative measure (37). This fact must be kept in mind when the subjects of investigation are selected from the superior intellectual levels.

3. Testing the Intelligence of Infants

When we turn to the question of the expected variability of test results with subjects of preschool age, the problem becomes very much more complicated. First, there is the increased difficulty of securing satisfactory coöperation on the part of children who are, as yet, little motivated by ideas of success or failure, although they may wish to please the examiner and are likely to participate in the test with considerable enthusiasm if the material happens to appeal to them. Second, there is the problem of the differential content of the tests devised for young children as related to that of tests used for older children, which makes it uncertain whether any lack of correlation between earlier and later tests results from basic idiosyncrasies in the pattern of growth or from incomparability of the material used for testing. Third, the small denominator used in computing the intelligence quotient when the chronological age is low brings it about that, unless the number of items available at each developmental level is materially increased (a thing that has been done in the more recent revision of the Stanford-Binet but that was not adequately taken care of in most of the earlier tests used), success or failure on a single item of the test will affect the total to a much greater extent than it does on a later test where the denominator of the fraction giving the IQ is larger.

Just what the performance on tests that have been devised for use with infants under the age of 18 months really signifies is unknown. That it is not entirely without meaning seems to be indicated by the fact that correlations between standing on repeated tests over a short period of time are usually positive and sometimes fairly high. That these tests of infants are not predictive of test performance at later ages has been shown so repeatedly that it is unnecessary to review the evidence here.

In this connection I might cite the results of a small and unpublished experiment of one of my graduate students, Miss Josephine Conger, which was carried out some years ago at a hospital for unmarried mothers in Minneapolis. Miss Conger gave the Linfert series of baby-tests to a group of 25 babies twice on successive days at each of the ages, one month, two months, and three months, making a total of 6 testings for each infant All the mothers were given both the Stanford-Binet and the Arthur Performance Scale. In addition, a considerable amount of information regarding the home background of the mothers and of the putative fathers, as well as of data on the physical status and birth histories of the infants, was secured. Although the tests given to the infants were consistent enough to make it evident that the individual differences could not be wholly due to chance, attempts to relate the developmental quotients to other factors yielded results that were discouragingly negative. With one exception, the correlations hovered persistently around zero. The single exception to this rule was the correlation between developmental quotient and length of time since the last feeding! All six of these correlations were positive and some were of fair magnitude. A naïve interpretation of these findings might lead to the conclusion that the way to make a baby intelligent is to starve him, but I doubt that we are as yet prepared to go quite to such lengths!

The bearing of these demonstrations of the unreliability and obscure meaning of tests of infants upon our main problem is perfectly obvious; much of the newer evidence on the effect of a change in environment upon child intelligence has been based upon just such early tests. About the only thing of which we are sure is this: regardless of whether or not a change in environment has taken place, changes in the intelligence quotients of young children are frequent and sometimes of astonishing magnitude if taken at their face value.

I might perhaps cite here the case of the little son of a scientist of international fame, who was first tested at our laboratory at the age of 22 months. He was not overtly shy or negativistic; on the contrary, he gave every indication of having a rollicking good time. But his performance on the test gave him an IQ of 86. However, there were plenty of evidences in his general behavior to indicate that this was not at all representative of what he could do if he had felt so disposed. The record was accordingly marked 'unreliable,' and the youngster was sent home with the suggestion that he return for a retest after a few weeks. On the next trial, about three months later, his IQ had jumped to 168 where, approximately, it has remained ever since.

4. The Hazards in Determining and Interpreting IQ Change

When the measure used is not IQ, but *IQ change*, and the first test is taken at an age at which instability rather than stability of test result is the rule, the hazards of interpretation are many times increased. Not only is the change in standing affected by two errors of measurement in place of one, but in any selected group these errors also have a disconcerting habit of taking on a systematic trend instead of being distributed at random. The effect of simple regression to which attention has repeatedly been called in this section is but one instance of this kind. The greater variability of high as compared with low IQ's that Terman and Merrill (37) have pointed out is another complicating factor when 'changes' from all levels of ability are combined in a single frequency table and treated by the simple rules of arithmetical averaging. Actually, the mathematical complications involved in the study of changes in individual IQ's are so many and varied that only a statistician of the first rank is equipped to deal with them.

All in all, the doctrine that the range of individual potentialities in the field of intelligence is so great that it may be shifted at will from one extreme to another is, in my judgment, a gross overstatement of the facts at hand. If, with Terman, we define intelligence as 'the ability to think in abstract terms,' it is but reasonable to suppose that *some* improvement in this ability will result from the acquisition of greater amounts of skill and knowledge that are the tools of thought. Other things being equal, better tools are an aid to good workmanship. But tools do not make the workman, and the extent to which the quality of work can be improved by the mere supplying of better implements is small in comparison with the differences between workers in their ability to make use of such ordinary facilities as are available to all.

Even in childhood, environment is not a thing apart from the individual; it is largely a matter of one's own making. Of two brothers reared in the same household, attending the same schools, with, as far as can be judged, equal opportunity to gain intellectual experience, one spends half his time at the library and most of his pocket-money on hobbies that contribute to some phase of useful knowledge or skill; the other is rarely to be seen with a book, and his pennies go for candy and motion pictures. Now there is little doubt that the ability of the first lad to score well on intelligence tests is facilitated to some extent by the increase in vocabulary that comes from his reading and the specialized skills and knowledge of mechanical principles or similar matters that he develops as a result of his intellectual interests. Perhaps, to that extent, we may say that his intelligence has been improved over what it would have been, had he been reared in an isolated mountain community where such things as libraries were non-existent, though it is but fair to note that, had this been the case, he might perhaps have acquired other types of skills and other items of knowledge that might be just as significant for his mental development, even though not included among those required to pass a standardized intelligence test.

The point has often been made, and quite properly, that even within the same family and the same household, the environment of the various members differs in very material respects. Those who stress the lability of intellectual progress are wont to use this argument as an explanation for the marked differences in intelligence so often observed, even among brothers and sisters. No one, so far as I am aware, has ever attempted to reduce these matters to an orderly system whereby it might be predicted that, certain conditions being true, this child would rank lower in intelligence than that one. The differential effect of these conditions is in reality purely a matter of conjecture; I might almost say that the entire concept is a means of escape to which the harassed investigator may always turn when facts become too stubborn. Granting, however, that such differences in intrafamily environment do exist, it may be well to give some thought to the contribution that each individual himself makes to the variance of his own environment before pinning too much faith upon the unproved hypothesis that the same kind and amount of variance induced by external factors will have the same effect — if, indeed, it is ever possible to bring such a thing about.

References

(1) Anderson, John A. "The prediction of terminal intelligence from infant and preschool tests." (Part I, Chapter XIII, of this *Yearbook*)

(2) Atkinson, M. C. "A study of infants of feebleminded delinquent females." *Jour. Psycho-Asthenics*, 32: 1926–1927, 110–128.

(3) Barrett, H. E., and Koch, H. L. "The effect of nursery-school training upon the mental-test performance of a group of orphanage children." *Ped. Sem. and Jour. Genet. Psychol.*, 37: 1930, 102–122.

(4) Burks, B. S. "The relative influence of nature and nurture upon mental development: A comparative study of foster parent-child resemblance and true parent-true child resemblance." [In] *The Twenty-Seventh Yearbook, Part I,* of this Society, 1928, pp. 219–316.

(5) Goodenough, F. L. "A preliminary report upon the effect of nursery school training upon the intelligence test scores of young children." [In] *The Twenty-Seventh Yearbook, Part I,* of this Society, 1928, pp. 261–369.

(6) Gordon, H. *Mental and Scholastic Tests among Retarded Children: An Inquiry into the Effects of Schooling on the Various Tests.* (Board of Education: London, 1923, Educational Pamphlets No. 44) 92 pp.

(7) Hildreth, G. "The effect of school environment upon Stanford-Binet tests of young children." [In] *The Twenty-Seventh Yearbook, Part I,* of this Society, 1928, pp. 355–359.

(8) Kawin, E., and Hoefer, C. *A Comparative Study of a Nursery-School versus a Non-Nursery-School Group.* (University of Chicago Press: Chicago, 1931) 52 pp.

(9) Kephart, N. C. "Influencing the rate of mental growth in retarded children." (Part II, Chapter XIII, of this *Yearbook*)

(10) KLINEBERG, O. *Negro Intelligence and Selective Migration.* (Columbia University Press: New York, 1935) 66 pp.

(11) LAWRENCE, E. M. " An investigation into the relation between intelligence and inheritance." *Brit. Jour. Psychol., Monog. Suppl.,* 16, No. 5: 1931, 80 pp.

(12) LEAHY, A. M. " A study of certain selective factors influencing prediction of the mental status of adopted children in nature-nurture research." *Ped. Sem. and Jour. Genet. Psychol.,* 41: 1932, 294–329.

(13) LEAHY, A. M. " Nature-nurture and intelligence." *Genet. Psychol. Monog.,* 17: 1935, 235–308.

(14) LITHAUER, D. B., and KLINEBERG, O. " A study of the variation in IQ of a group of dependent children in institution and foster homes." *Ped. Sem. and Jour. Genet. Psychol.,* 42: 1933, 236–242.

(15) MCNEMAR, Q. " A critical examination of the University of Iowa studies of environmental influences upon the IQ." (Accepted for publication in the *Psychological Bulletin*)

(16) MALLER, J. B. " Vital indices and their relation to psychological and social factors." *Human Biol.,* 5: 1933, 94–121.

(17) NATIONAL SOCIETY FOR THE STUDY OF EDUCATION, *The Twenty-Seventh Yearbook. Nature and Nurture. Part I. Their Influence upon Intelligence. Part II. Their Influence upon Achievement,* 1928; ix. 465; xv, 397 pp.

(18) PETERSON, T. J. " A preliminary study of the effects of previous nursery-school attendance upon five-year-old children entering kindergartens." *Studies in Preschool Education.* (University of Iowa: Iowa City, 1937) pp. 197–248. (University of Iowa Studies in Child Welfare, Vol. 13, No. 4.)

(19) REYMERT, M. L., and HINTON, R. T. " The effect of a change to a relatively superior environment on the IQ's of one hundred children." (Part II, Chapter XVII, of this *Yearbook*)

(20) SEAGOE, M. V. " The transient child." *Jour. Juvenile Res.,* 16: 1932, 251–257.

(21) SHERMAN, M., and HENRY, T. R. *Hollow Folk.* (Thomas Y. Crowell and Co.: New York, 1933) 215 pp.

(22) SHERMAN, M., and KEY, C. B. " The intelligence of isolated mountain children." *Child Develop.,* 3: 1932, 279–290.

(23) SIMPSON, B. R. " The wandering IQ: Is it time for it to settle down? " *Jour. Psychol.,* 7: 1939, 351–367.

(24) SKEELS, H. M. " Mental development of children in foster homes." *Ped. Sem. and Jour. Genet. Psychol.,* 49: 1936, 91–106.

(25) SKEELS, H. M. " Mental development of children in foster homes." *Jour. Consulting Psychol.,* 2: 1938, 33–43.

(26) SKEELS, H. M. " Some Iowa studies on the mental growth of children in relation to differentials in environment." (Part II, Chapter XX, of this *Yearbook*)

(27) SKEELS, H. M., and FILLMORE, E. A. " Mental development of children from underprivileged homes." *Ped. Sem. and Jour. Genet. Psychol.,* 50: 1937, 427–439.

(28) SKEELS, H. M., UPDEGRAFF, R., WELLMAN, B. L., and WILLIAMS, H. M. *A Study of Environmental Stimulation: An Orphanage Preschool Project.* (University of Iowa Studies in Child Welfare, 15: No. 4, 1938) 191 pp.

(29) SKODAK, M. "The mental development of adopted children whose true mothers are feeble-minded." *Child Develop.*, 9: 1938, 303–308.

(30) SKODAK, M. *Children in Foster Homes: A Study of Mental Development.* (University of Iowa Studies in Child Welfare, 16: No. 1, 1939) 156 pp.

(31) SNYGG, D. "The relation between the intelligence of mothers and of their children living in foster homes." *Ped. Sem. and Jour. Genet. Psychol.*, 52: 1938, 401–406.

(32) SPEER, G. "The mental development of children of feebleminded and normal mothers." (Part II, Chapter XXI, of this *Yearbook*)

(33) STARKWEATHER, E., and ROBERTS, K. "IQ changes occurring during nursery-school attendance at the Merrill-Palmer School." (Part II, Chapter XXII, of this *Yearbook*)

(34) STIPPICH, M. L. "The mental development of children of feebleminded mothers." (Part II, Chapter XXIII, of this *Yearbook*)

(35) STUTSMAN, R. *Mental Measurement of Preschool Children.* (World Book Co.: Yonkers, New York, 1931) 368 pp.

(36) TERMAN, L. M. *The Intelligence of School Children.* (Houghton Mifflin Co.: Boston, 1919) 317 pp.

(37) TERMAN, L. M., and MERRILL, M. A. *Measuring Intelligence: A Guide to the Administration of the New Revised Stanford-Binet Tests of Intelligence.* (Houghton Mifflin Co.: Boston, 1937) 461 pp. (esp. pp. 44–45)

(38) THOMSON, G. H. "A formula to correct for the effect of errors of measurement on the correlation of initial values with gains." *Jour. Exper. Psychol.*, 7: 1924, 321–324.

(39) THORNDIKE, E. L. "The influence of the chance imperfections of measures upon the relation of initial score to gain or loss." *Jour. Exper. Psychol.*, 7: 1924, 225–232.

(40) THORNDIKE, R. L. "Retest changes in the IQ in certain superior schools." (Part II, Chapter XXIV, of this *Yearbook*)

(41) WELLMAN, B. L. "Some new bases for interpretation of the IQ." *Ped. Sem. and Jour. Genet. Psychol.*, 41: 1932, 116–126.

(42) WELLMAN, B. L. "The effect of preschool attendance upon the IQ." *Jour. Exper. Educ.*, 1: 1932–1933, 48–69.

(43) WELLMAN, B. L. "Growth in intelligence under differing school environments." *Jour. Exper. Educ.*, 3: 1934–1935, 59–83.

(44) WELLMAN, B. L. "Mental growth from preschool to college." *Jour. Exper. Educ.*, 6: 1937–1938, 127–138.

(45) WELLMAN, B. L. [Abstract of] T. J. Peterson. "A preliminary study of the effects of previous nursery-school attendance upon five-year-old children entering kindergarten." University of Iowa Studies in Child Welfare, 14, 197–248. Abstract No. 4436, *Psychological Abstracts,* 12: 1938, 492.

(46) WELLMAN, B. L. "Our changing concept of intelligence." *Jour. Consulting Psychol.*, 2: 1938, 97–107.

(47) WELLMAN, B. L. *The Intelligence of Preschool Children as Measured by the Merrill-Palmer Scale of Performance Tests.* (University of Iowa Studies in Child Welfare, 15, No. 3: 1938) 150 pp.

(48) WELLMAN, B. L. "The IQ: A reply." *Jour. Psychol.*, 8: 1939, 143–155.

(49) WELLS, J., and ARTHUR, G. "Effect of foster-home placement on the intel-

ligence ratings of children of feebleminded parents." *Mental Hygiene,* 23: 1939, 277–285.

(50) WOOLLEY, H. T. "The validity of standards of mental measurement in young childhood." *School and Soc.,* 21: 1925, 476–482.

(51) YULE, G. U., and KENDALL, M. G. *Introduction to the Theory of Statistics.* (Charles Griffin: London, 1937) 570 pp. (esp. p. 70)

CHAPTER XII

SOME SPECIAL PROBLEMS OF NATURE–NURTURE RESEARCH

FLORENCE L. GOODENOUGH
Research Professor, Institute of Child Welfare
University of Minnesota
Minneapolis, Minnesota

I. METHODS OF APPROACH

During recent years there has been a tendency to extol the so-called 'longitudinal' approach to the study of human development, in which the same group of children are followed for a period of years, in contrast to the 'cross-sectional' approach, in which comparisons are made between different groups of children of different ages, often of successive ages. There can be no doubt that the former method lends itself to the study of many problems for which the latter is poorly suited, that by using the same group of children from year to year, certain types of sampling errors are automatically excluded, and, what is of chief importance for our present consideration, it becomes possible to employ techniques that are truly experimental in nature, since the effect of an artificially interpolated condition can be studied in terms of responses both before and after its occurrence. Thus, if the group studied is of fair size, it becomes theoretically possible to make use of the techniques of investigation that have been devised for cross-sectional studies, and to supplement these by other methods that cannot readily be employed in such studies. But unmixed blessings are rare. Cross-sectional studies, it is true, are peculiarly subject to certain hazards and limitations; yet the longitudinal approach is likewise subject to hazards and to fallacies more or less peculiar to itself. So many of the studies we wish to consider are of the longitudinal type that it seems pertinent to point out some of the more important of the hazards likely to be encountered when this method is employed.

II. SAMPLING ERRORS

Because of the difficulties of maintaining contact with children over any considerable period of time, selective factors that affect the character of the group being studied almost inevitably occur in any long-time investigation. Conclusions therefore must be limited to groups similar to the one studied. Ordinarily these will be groups of relatively superior ability and economic stability because, even when no coöperation on the part of parents is required (as would be the case when all measurements are made in the public schools and there is no direct contact with the home), there is commonly some selective effect due to the greater lability of domicile among persons of lower economic status than among those of superior status. If, as is true in practically all investigations in which young children are used, there is the further requirement of active coöperation on the part of the parents in bringing the children for measurements, supplying information about their development, and so forth, then the selective effect just mentioned becomes much greater. All this need not affect the validity of any study, provided sufficient information is given to make clear what factors were operative both in determining the original selection of subjects and also in retaining them as members of the group during the time it was studied.

On the other hand, the danger that the sampling of subjects at the successive ages will be non-representative in respect to factors other than age, either as a result of chance variation or of selective factors that operate unequally for children of different ages — one of the main drawbacks of the cross-sectional method — is automatically eliminated by the longitudinal method, and it is from this fact that it derives one of its main claims to superiority for the study of mental and physical growth. It follows that the longitudinal method requires fewer cases for the establishment of growth principles, since the developmental patterns are not confused by irregular distribution at successive ages of individual differences in respect to the trait studied.

This has been shown graphically by Shuttleworth [1] and is demonstrated statistically by the principle that a correlation between measures decreases the size of the difference between them and thus increases the critical ratio to a point at which it is commonly regarded as 'statistically valid' when the number of cases is far smaller than would be needed to establish the existence of the difference in question with an equal degree of assurance if the measures were

[1] F. K. Shuttleworth. *Sexual Maturation and the Physical Growth of Girls Age Six to Nineteen.* (Monographs, Society for Research in Child Development, Vol. II, No. 5. National Research Council: Washington, D. C., 1937) 252 pp. (esp. Chap. 3)

uncorrelated. Since individual differences in most traits that have significance in the life of a child tend to persist over a period of years, some degree of correlation between earlier and later standing is the rule.

It is, of course, taken for granted in our discussion of the longitudinal approach that we are considering only truly longitudinal studies, in which there have been no changes in the group of cases studied; more exactly, that such changes in group composition as inevitably will occur in the course of a long-time study are eliminated by excluding from each comparison all cases not included in *both* of the measurements to be compared. Thus, if 50 children are measured at the age of 5, and only 40 of them are available for remeasurement at the age of 6, the 10 who have dropped out during the year must be omitted in computing the initial as well as the final mean. Otherwise, the sampling error is likely to be greater in a 'longitudinal' than in a 'cross-sectional' study.

But even the precaution just mentioned may not completely take care of sampling errors due to selective elimination from the group if, as is very likely to be the case, the errors of measurement are not distributed at random but themselves show a systematic trend, so that the obtained measures are more likely to err in one direction than the other, and the elmination is on the basis of the true scores rather than on the basis of the obtained scores. If, for example, young children are tested by examiners who, because of inexperience or ineptness in handling them, frequently fail to elicit the best responses that the children are capable of giving (with the result that the test results are more likely to be too low than too high), and if, furthermore, the children who are truly bright are more likely to be left in the group than those who are truly backward, then a systematic error will appear when the findings from early and later measurement are compared. For if, at later measurement, the skewness due to errors of measurement tends to be corrected (as would be the case if such factors as shyness, unfamiliarity with the examiner and the testing situation were largely responsible), there will obviously occur an increase in mean score that does not reflect a true increase in ability, but only an improvement in the systematic error of measurement. Moreover, when two groups are compared, if the tendency toward selective elimination on the basis of true status is greater in one group than in the other (that is, if the correlation between true intelligence and the chances of being eliminated is higher, let us say, for a control group than for an experimental group), then it is to be expected that the apparent increase in status will be greater for that group in which a higher proportion of the truly backward have dropped out between initial test and retest.

It is necessary, therefore, to look into the question of differential selection with much care before assuming that a difference in the measured change in status of two groups is due to the experimental situation being studied rather than to differential factors affecting elimination of subjects from the study or to differential effects of the experimental situation upon errors of measurement.

We need not comment on the fact that the presence of a systematic error of measurement will play an even more important rôle in producing differential changes in the measured status of an experimental and a control group if one effect of the experimental situation is to minimize or do away with a systematic error of measurement originally present in both groups. If, for example, in the situation just described, the systematic error was due to shyness and if the experimental situation, say nursery-school training, caused children in a nursery-school group to become less shy, whereas the children of the control group showed no change in this respect, a difference in the amount of measured gain from initial to later testings would inevitably occur. In such a case the real change is in the error of measurement; the true score is not necessarily affected at all, in spite of what may be a very considerable change in the obtained score.

III. The Use of Control Groups

Because of the large number of factors that may either change the form of the mental-growth curve or give a spurious appearance of such change, it is hazardous to assume that an apparent modification of the course of growth of a selected group, ensuant to the interpolation of some known factor, has real existence, or that, if real, it is the result of the factor under consideration and not of some other undetermined variable, until all possible checks have been made to ascertain the facts. It is well for both the experimenter and the reader of experimental data not to lose sight of the time-honored principle that in any controversy *the burden of proof rests with him who makes the claim.* Although it is never possible to prove a universal negative, it is possible, in any specific instance, to state whether or not the evidence submitted warrants the conclusions drawn.

Recognition of the factors just mentioned has led to the widespread adoption of a number of techniques designed to bring a greater number of the possible variables under experimental control. Among these, the use of *control groups,* selected to resemble the experimental group in as many factors as possible, is one of the best known.

It is important to note that the control-group technique has two forms that are of unequal dependability. The first is that of the experimental laboratory in which the experimenter selects from a total population those cases that are to constitute the experimental group and those that are to be used as controls. The method of selection depends upon the complexity of the problem to be studied and the number of subjects to be used. If the latter is fairly large and the problem is not highly complicated by reason of correlation with many other factors, then some system of random selection (such as choosing alternate names for each group) is to be preferred, since it eliminates all possibility of differences between the groups except those due to chance.

If the number of cases is small or the number of correlated variables large, then the method of *matched pairs* is commonly used. The chief difficulty, of course, is to include all significant variables in the matching, and unless the population from which the two groups to be selected is rather large or the number of correlated variables small, the matching is at best likely to be only approximate. It should be noted, too, that when critical ratios are computed for paired groups, account must be taken of the correlation between the pairs brought about by the matching.

Even when the problem of matching is thus within the control of the experimenter, the possibility of significant differences between the paired groups is not completely excluded. Moreover, when the second of the two possible conditions obtains — when the composition of the experimental group is predetermined and the experimenter is forced to match his control subjects to this group of selected subjects as best he can — then the difficulty of ensuring two comparable groups is greatly increased.

Consider, for example, the studies on the effect of nursery-school training presented in Part II of this Yearbook. Only a small proportion of children of nursery-school age actually attend a nursery school. It is not unreasonable to suppose that attendance or non-attendance is due to a large number of factors that will vary with the community, the school, and the conditions of entrance, as, for example, whether the school is free or tuitional, the amount and kind of publicity that is given to it, its location, transportational facilities, amount of coöperation expected from parents, and other more subtle factors. Rarely are all these factors known to the investigator, and even when known or suspected, it frequently happens that no dependable data are available for matching the control to the experimental group. Lamson's study (in Part II) is a case in point. She found that of the traits measured by her, only one showed a significant difference between the nursery-school and the control group. According to the ratings of kindergarten and primary-grade teachers, children who had previously attended a nursery school showed distinctly poorer personality adjustment than did those who were without nursery-school training. Yet we are scarcely warranted in assuming that the nursery school in question had an unfavorable effect upon child personality. First of all we should need to ask whether there was a greater likelihood that children showing undesirable personality traits would be sent to the nursery school than others of more pleasing characteristics, or whether the parents who were particularly inept at child training were more likely than others to shift the task to other hands. This might readily be the case if, through popular lectures or other means, the parents had been led to feel that attendance at a nursery school is an effective means of correcting undesirable behavior in young children. If it so happened that the nursery-school population was made up in large part of problem children or, what is much the same thing, the children of problem parents, it is unfair to interpret the later superiority of Lamson's control group as evidence against the nursery school. It may only show that the school was faced with a problem greater than it was equipped to solve. For it cannot be too persistently kept in mind that to establish a 'statistically significant' difference between groups does not in itself tell what the significant difference is; that must be determined by other evidence.

It may be said that if Lamson had been able to secure evidence about the personality of her subjects at the time of nursery-school entrance, this matter would have been clarified. That is true in a measure, but the problem is too complex to be answered by a simple 'before and after' treatment. Let me illustrate by a hypothetical case.

Suppose that in a certain community the idea that shyness in children is a serious fault had been greatly stressed, along with the notion that nursery-school attendance is a good way to overcome the fault. Almost certainly shy children would be more likely than friendly children to be placed in nursery schools in that community. But shyness does not operate in a vacuum; it affects the child's reactions along many lines, including his performance on standardized mental tests, wherein his rating will inevitably be lowered. Now if, with maturing and the social experience gained in the nursery school, a gain in self-confidence is made, the child's test score will tend to revert to what it would originally have been had the measure been adequate; but that obviously would not warrant the conclusion that his intelligence had increased.

It is in cases such as this that even the apparently well-controlled study with experimental and control groups well matched as to measured initial ability is likely to break down and lead to misinterpretation of the quantitative results.

The specific example given is but one of many that may operate in a similar fashion. The reader should note, also, that in such instances the experimental group will not only show a seeming gain but also a gain greater than that of the control group, with an amount of difference possibly great enough to meet the criterion for statistical significance. What the gain signifies, however, may be differential selection at the outset of the experiment and not necessarily improvement in standing as a result of nursery-school attendance.

An increase in test score may come about either by increasing the final score or by lowering the initial score. Both possibilities must be examined in interpreting the increase.

This is nicely illustrated in a study by Schott [1] on the effect of placement in foster homes upon the intelligence quotients of children. Children were tested at the time of placement and retested about a year later, at which time a mean gain in intelligence quotient of about 4 points was found. Owing to an error in computation, Schott stated that this difference was not statistically significant. He also pointed out that the small gain was perhaps attributable to a spurious lowering of the initial score as a result of giving the test at a time when the children were under considerable emotional stress as a result of having just been removed from their own homes and placed in various temporary homes pending final disposition. Although it was impossible to determine which of the two possible explanations — actual increase in the final score as a result of the improved environment or spurious lowering of the initial score as a result of the emotional situation at the time of the first test — was the correct one, Schott presented further evidence that makes the latter hypothesis appear the more tenable.

However, in a brief note published somewhat later, Thorndike called attention to the computational error in Schott's article and pointed out that, when this error was corrected, the difference of 4 IQ points became 'statistically significant.' He then proposed to correct Schott's conclusion so that it would read: "after a year's residence in a foster home, a statistically significant increase in intelligence quotient was found," but he did not mention the possibility — if not the probability — that this obtained difference in test score might 'signify' something very different from that which would be implied by his proposed corrected statement.

---

[1] E. L. Schott. "IQ changes in foster-home children." *Jour. Applied Psychol.,* 21: 1937, 107–112.

R. L. Thorndike. "Note on 'IQ change in foster-home children,' by Emmett L. Schott." *Jour. Applied Psychol.,* 22: 1938, 662.

F. L. Goodenough. "A comment on Robert L. Thorndike's 'Note on IQ changes in foster-home children,' by Emmett L. Schott." *Jour. Applied Psychol.,* 23: 1939, 308–309.

IV. The Possibility of Systematic Errors in Test Administration

No one acquainted with the experimental literature on suggestion or familiar with the routine precautions taken in all psychological laboratories to avoid systematic errors due to prejudgment of stimulus or to knowledge of the expected nature of the response can fail to be cognizant of the extent to which similar factors may sway the results of experiments in which individual mental tests are used.

For it is at once the chief advantage and the greatest hazard of the individual test, as compared to the group test, that the former provides for certain variations in procedure that are impossible when the latter is used. In the individual test, a question misunderstood by the child may as a rule be repeated, sometimes in another prescribed phrasing. The rate of questioning is adapted to the individual child; if he hesitates for a reply, the examiner is in most instances free to use his own judgment in deciding how much, if any, urging is to be given, what form such urging shall take, and how long to wait before passing on to the next item. Moreover, the very fact that the child and the examiner are associated but by themselves makes the nature of the personal relationship between them important in determining the test results. Especially is this true of little children who are so greatly motivated by a desire to please the examiner and to win his commendation. Since children differ greatly in speed of response, attitude toward praise, preference or distaste for certain kinds of tasks, and the like, it is obvious that the flexibility of procedure in handling the child permitted by an individual test obviously gives it advantages over the group test in which all children are treated in exactly the same manner.

The skill of the examiner undoubtedly plays a far greater part in determining results in individual testing than it does in group testing. Fifteen minutes' instruction is commonly enough to teach almost any competent graduate student to give a group mental test. In contrast, from 100 to 200 hours of instruction, demonstration, and practice may well be needed to insure equal proficiency in individual testing; indeed, there are a fair number of otherwise capable students who never do acquire this skill. For the safe interpretation of any report based on the results of individual tests, precise information about the training and experience of the examiners is, therefore, essential, and the younger the subjects tested, the more important does this information become.

But even with thoroughly competent examiners, it is highly unsafe to assume that test results will be free from systematic errors if the examiner at the outset of the test has information that leads him to anticipate the results. Then the flexibility of procedure that is otherwise an asset becomes a distinct hazard, for it requires only minor variations in those aspects of the procedure that the examiner is free to vary to produce perceptible changes in test results.

For example, the examiner who anticipates that a particular child will in any case do poorly on the test is likely, without himself being aware that he is doing so, to be much more ready to accept an " I don't know " at its face value, to shorten the period of waiting for a reply when the child remains silent, or even to abbreviate the range of testing beyond that demanded by the rules, by assuming that it is " not worth while " to try a particular task since (in the examiner's opinion) the child would certainly be unable to pass it. On the other hand, if the presumption is that the child is bright, there is equal likelihood of an unconscious attempt on the part of the examiner to secure the maximal performance of which the subject is capable. The amount of urging may be doubled or trebled, praise greatly increased, and every interest-arousing trick in the examiner's repertoire may be brought into play in order to secure the expected result. Not that the differences here described are necessarily or usually the result of conscious intent: they are merely the outcome of human nature behaving naturally. Few of us are so heroic as to exert maximal effort to secure a result that we are convinced, at the outset, is impossible, and most of us like to see our expectations confirmed. Safety from this bias is only to be found in the complete absence of prejudgment.

It is unfortunate that no objective experiment has ever been made, as far as I am aware, of the effect upon the correlation between tests of preknowledge, on the part of the examiners, of the results of previous testings of the same subjects. Without definite knowledge of the extent to which this factor can affect results, I can only say that I do not trust myself to be uninfluenced by it. Safety is only to be found in ignorance.

V. Systematic Errors in Test Scoring

In many items in the Binet scale, the scoring of ' borderline ' responses presents a problem, even to experienced examiners. No matter how long one tests, new responses, for the scoring of which there is no absolutely fixed rule, continue to crop up. Here, obviously, is another opportunity for systematic errors to arise, and the extent to which they may do so in individual instances has never been adequately determined.

In this connection I may cite a small unpublished experiment of my own. At that time I was conducting some intensive experiments with the children of a fifth-grade class in one of the eastern cities and had occasion to examine the results of a spelling test that had been given and scored by the classroom teacher. I found a fairly large number of errors in scoring that suggested a systematic trend. To check this hypothesis, I prepared a rating scale made up of a number of items purporting to measure 'personal attractiveness.' By intent, the ratings were so phrased as to yield as subjective an estimate as possible, in the hope of learning something about the teacher's personal attitude toward the individual children. When the summed ratings were correlated with the number and direction of the errors in scoring, an r of approximately +.40 was obtained. Errors by the children whom the teacher judged most attractive had frequently been overlooked; correct spellings, if at all illegibly written by the less-favored members of the class, were in many cases marked wrong. The teacher was, of course, entirely innocent of any intentional unfairness and was completely ignorant of any irregularities in her grading. But human nature plays strange tricks with all of us.

VI. Other Factors Affecting Test Scores: Specialized Experience

Elsewhere[1] I have pointed out that a mental test is a sampling device rather than a measuring device in the proper sense of the term and that its validity therefore depends upon the principles of sampling, over and above whatever principles of measurement may apply to its separate items. The items of which a test is composed theoretically constitute a carefully selected sampling of the abilities of the individual to whom the test is given, and his performance on these items is assumed to be representative of his probable performance on other items not included in the sample. It is apparent, therefore, that if, as a result of some kind of specialized training or experience, a child's ability to perform the particular set of items included in the test is improved, without corresponding change in his ability to cope with the vast number of other items of the total life-situation, little of real value has been gained (since the test items constitute so infinitesimal a proportion of the total), but the diagnostic value of the test itself has been greatly diminished, or even completely lost, since the conditions requisite for a valid sampling are no longer fulfilled. 'Specialized training,' in the sense that the term is here used, is not limited to

[1] F. L. Goodenough. "The measurement of mental functions in primitive groups." *American Anthropologist,* 38: 1936, 1-11.

the obvious cases of intentional coaching on the test; it includes also any kind of experience in which these items or others closely related thereto are emphasized to a greater degree than those not included in the test sample. Hence, we prefer to use the term 'intelligence-test score' rather than its presumed correlative 'intelligence,' because the conditions under which the two may be assumed to parallel each other are so readily disturbed.

The effect of specialized training may show itself in a wide variety of ways. Suppose, for example, that the nursery-school training given the children in Lamson's experimental group was of a kind that emphasized much freedom of activity, but that the program of the kindergarten and primary grades was comparatively formalized. It might well be the case that the children who had previously come to associate 'school' with a good deal of active romping, comparative freedom of choice, and absence of restrictions might actually find it harder to adjust to a very different type of school régime than would those who had no previously acquired habits or attitudes with respect to school behavior. Again, it might well be that the games, activities, and play materials of the nursery school were so similar to those used later that the children had tired of them, and thus the later school experience failed to stimulate their interest. In either case it would be fair to say that the experience in the nursery school in question had an unfavorable effect upon adjustment to the kind of school régime later encountered, but it would be hazardous to assume that the unfavorable effect extended to situations outside the schoolroom. The situation is not unlike that found in the applications of factor analysis, where it is relatively easy to isolate one or more 'factors,' but very difficult to determine their nature. It is easy to ascertain that a significant difference exists, but not at all easy to find out what it signifies.

VII. The Question of Test Comparability

The amount of available information concerning the comparability of the various tests and scales used for measuring intelligence is at best scanty, and particularly so at the early ages. In the Minnesota study[1] on the effect of nursery-school training, for instance, it was found that the intelligence quotients earned on the initial Minnesota tests were somewhat lower, at all ages, for all occupational classes, and for both nursery-school and control groups, than those earned by the

[1] See Part II, Chapter IX.

same children on either the 1916 or the 1937 Revision of the Stanford-Binet at later ages. This may indicate that the Minnesota norms are slightly higher than those for the Stanford, but the fact that the difference steadily increases with advancing age at the time of the Stanford test — a fact not consonant with the tendency reported for Terman's standardization group, where there is a small but fairly consistent trend toward a drop in mean IQ for the successive ages (a trend that has been verified by a number of other investigators since the publication of the scale) — suggests that some other factor may be operative. However, even if such a difference exists, it is possible to make allowance for it when all the scores reported in any single distribution are based on the same test.

Because there is some similarity in the content of the Stanford and that of the Kuhlmann Revisions of the Binet, a number of investigators have followed the practice of using IQ's from the two revisions interchangeably within the same distribution. Before such a practice can be justified, two things must be demonstrated to be true: first, that the scales are psychologically equivalent in content at all age levels, and second, that they are statistically equal in difficulty and in variational tendency at each age. If the first of these conditions is not fulfilled, there is nothing that can be done about it, but if the discrepancies are solely of the second type, it is possible to transmute the scores earned on one scale into their equivalent values on the other scale without too great loss of accuracy, provided, of course, that the table of equivalents is based upon a sufficiently large number of comparable scores to yield stable values. As far as I am aware, no adequate study of the equivalence of the two tests at the early age levels has appeared to date, but it may be noted that the irregularities at the early age levels of the Kuhlmann-Binet that I have reported elsewhere [1] make it appear very unlikely that results from the two measures can be used interchangeably without distortion of the facts.

VIII. The Question of the Psychological Equivalence of Test-Content at Widely Separated Age Levels

Repeated investigations have shown that few, if any, of the tests used to measure the developmental status of infants have predictive

[1] F. L. Goodenough. *The Kuhlmann-Binet Tests for Children of Preschool Age: A Critical Study and Evaluation.* (University of Minnesota, Institute of Child Welfare Series No. 2: University of Minnesota Press, 1928) 146 pp.

value for later mental status.[1] A number of theoretical explanations may be advanced for this. The first is that the rate of individual mental growth is not constant, but that the established 'constancy of the IQ' is wholly an artifact resulting from the overlap between initial and final status. It must not be forgotten that a mental test is not a measure of rate, but of status, and that in so far as an IQ provides a crude approximation to relative rate of growth (compared with that of the average child), it is always an *average* rate, based upon total mental progress up to the time the test is given. An IQ thus includes not simply a measurement of growth accomplished since any preceding tests were given, but also a measurement of that accomplished previous to the administration of the earlier tests. Hence, even if there is no 'constancy' whatever in the *rate* of growth, a correlation between the two measurements will inevitably result[2] because of the common element. The results of the two tests may accordingly be expressed as follows:

$$\text{Initial test score} = S_1 + E_1$$
$$\text{Final test score} = S_1 + I + E_2$$
$$\begin{aligned} \text{Where } S_1 &= \text{true mental status at time of first test,} \\ I &= \text{true growth increment between tests, and} \\ E_1 \text{ and } E_2 &= \text{errors of measurement.} \end{aligned}$$

These 'errors' include not only random errors, but also any systematic errors not common to both tests and any discrepancies resulting from 'psychological' differences in the content of the items making up the test at the two age levels. Now if the combined value of E_1 and E_2 is very great, the importance of S_1 may be partially or, in some cases, entirely masked, so that the correlation between earlier and later measurement will approach zero. It is obvious that this is most likely to occur when the value of S_1 is small; that is, at the early ages, or when the two measurements are psychologically dissimilar in character. The effect of *I* upon the correlation between earlier and later testings will depend upon its absolute magnitude, which will vary with the interval between testings, and upon the magnitude of

[1] P. H. Furfey and J. Muehlenbein. "The validity of infant intelligence tests." *Ped. Sem. and Jour. Genet. Psychol.*, 40: 1932, 219–223.

N. Bayley. "Mental growth during the first three years: A developmental study of sixty-one children by repeated tests." *Genet. Psychol. Monog.*, 14: 1933, 1–92.

[2] Provided, of course, that the test is at least to some extent a valid instrument of measurement.

its correlation with S_1.[1] If, as some have too hastily taken for granted, this correlation is +1.00, then the effect of *I* upon the correlation is zero. If, however, as recent studies indicate is more likely to be the case, the correlation is considerably less than +1.00, which means that there are true variations in the rate of individual mental growth, then the variations in the *I* factor may be great enough to lower the correlations between tests very materially, especially if S_1 is small.

Anderson [2] has given an excellent theoretical discussion of this principle, and has drawn up a series of curves in which the hypothetical values are compared with those derived from the empirical data of a number of published studies. It is apparent from his analysis that, if the proportion of the total score that is included in S_1 is large, correlations of a magnitude usually regarded as 'high' may be obtained between test and retest in spite of very material changes in growth rate during the interim between testings; again, that these correlations may be very low, even though the true rate of growth be extremely even from age to age, if S_1 is small and the error of measurement large. It follows, therefore, that much of the so-called greater 'plasticity' of the young child is nothing more than a mathematical necessity of the time factor. Suppose, for example, that there is no increase in test score over a period of one year — a condition that may appear either as the result of true arrested or retarded development or from errors of measurement: in the case of a child of 14 with an initial IQ of 100, his IQ would fall to 93 at the end of the year; but in the case of a child of 1 year who failed to gain in absolute score for one year his original IQ of 100 would fall to 50, which seems like a very startling change indeed. Actually such a change occurring in the measured IQ is much less disturbing than it seems when it occurs at so early an age. The significant point is the small size of the common factor, S_1, which means that both errors of measurement and small changes in the expected size of the true growth increment will produce large changes in the intelligence quotient from test to test. As age at initial measurement

---

[1] This statement, of course, refers to the correlation between true scores and not between measured scores, which include the errors of measurement previously mentioned.

[2] J. E. Anderson. "The limitations of infant and preschool tests in the measurement of intelligence." *Jour. Psychol.*, 8: 1939, 351–379. See also Part I, Chapter XIII.

increases, changes in the increment become of less and less significance in determining the mental level, just as a final burst of speed will accomplish little for the runner who has already fallen too far behind in the race. With advancing age either the intelligence quotient or any other derived measure that is expressed in relative terms inevitably becomes a better indicator of later status than it was at an earlier age because of the increasing magnitude of the common factor, S_1. The only condition under which the predictive value would not change with age would be that in which the correlation between S_1 and I is $+1.00$ and the error of measurement is proportional to the age at measurement.

IX. The Measurement of Gain or Loss

It has already been pointed out that the dependability of a measured gain or loss is affected by errors of measurement in both the initial and the final score. 'Gain' is of necessity a very unstable measure unless the basic measurements from which it is derived are exceptionally accurate. If systematic, as well as random, errors of measurement are present, the interpretation of a measured gain becomes still more hazardous. Further difficulties arise when an original group is broken up into subgroups on the basis of special factors of various kinds, unless it can be demonstrated that the basis of selection is unrelated to the size or the direction of the experimental errors involved.

One example of this kind continually recurs in the experimental literature; namely, the discovery that groups originally measuring low will apparently 'gain' more upon retest than groups originally ranking higher, and vice versa. This is, of course, a necessary effect of unreliability of measurement, since a selection of those cases whose original scores were low will inevitably include an undue proportion of those whose scores (due to errors of measurement) were *too low*. Upon remeasurement there will always be a tendency for individual scores to regress toward their true means, since random errors of measurement are uncorrelated with each other. It follows, therefore, that when groups that include an excess proportion of 'downward' errors of measurement are retested, the mean score will always be increased on the second measurement, and while it is, of course, entirely possible that other factors may operate to bring about an additional increase that is real, or to lower the result so that the change due to regression may be cancelled, the extent of the regressive tendency must always be determined and allowed for before it is safe to assume that additional factors have actually been operative. If, and only if, the obtained change is greater than that due to regression, need other factors be sought. Thomson [1] has derived a formula by which such correction can be made, and a simplified expression of this formula has recently been developed by Zieve,[2] who presents a table showing the wide variations in the obtained correlation between initial scores and gains that result from varying degrees of unreliability of measurement when the true correlation remains the same. It is unfortunate that this formula has not been more widely used, since many questionable interpretations of the results of retests might then have been avoided. As an example, a number of studies in which the so-called 'accomplishment ratio' [3] has been used as a means of comparing accomplishment with ability have reported that bright children are less likely than backward ones to 'work up to their ability.' This statement may or may not be true, but its truth cannot be demonstrated by a comparison of the mean accomplishment ratios of the two groups unless correction for regression is made. A further example is to be found in certain studies in Part II of this Yearbook and elsewhere that stress the greater improvement after a period of nursery-school training of children whose original test scores were low. A number of possible explanations for this discrepancy have been proposed, upon the assumption that the obtained differences indicate a true differential effect of training upon the several groups. If the regression due to errors of measurement be taken into account, it is likely that none of these explanations would be necessary, since examination of the data makes it appear very probable that regression is the sole factor at work. It is, however, impossible to check this hypothesis from the published data.

X. Summary

The statistical pitfalls inherent in the longitudinal approach to the study of mental growth are no fewer in number than those encountered in cross-sectional studies, though they differ somewhat in kind. Because the longitudinal method has been less widely used, these sources of error have been less generally understood or guarded against. In this chapter, some of the more common of these sources of error

[1] G. H. Thomson. "A formula to correct for the effect of errors of measurement on the correlation of initial values with gains." *Jour. Exper. Psychol.*, 7: 1925, 321–324.

[2] L. Zieve. "Note on the correlation of initial scores and gains." (To appear in *Jour. Educ. Psychol.*)

[3] The formula for the accomplishment ratio is as follows:

$$AR = \frac{EA}{MA} \text{ or } \frac{EQ}{IQ}$$

where EA and EQ are based upon educational tests and MA and IQ upon intelligence tests.

in procedure or in interpretation of results have been pointed out, together with suggestions for bringing them under better experimental control in future studies. The errors considered include:

1. Errors brought about by selective elimination of many of the original subjects during the course of the study, as a result of which those individuals who remain are unlikely to constitute a representative sample of the initial group, either in respect to status as first measured or, a fact of more consequence when 'constant' groups are compared, in respect to whatever potentialities of growth may be tied in with the chances of elimination.
2. Errors resulting from unequal matching of experimental and control groups when the selection of the former is not within the control of the experimenter, or from unequal experience of the two groups in respect to matters affecting the results of the measurement used for comparison, but not affecting the trait that it is presumed to measure (uncontrolled variables).
3. Systematic errors of measurement due to prejudgment of results when examiners are acquainted with results of preceding measurements or anticipate that findings will tend to take a given direction. Such errors may affect either the administration of tests or their scoring, or both.
4. Errors resulting from the non-comparability of tests used at different ages, particularly when the time span between initial test and retest covers the interval from early to later childhood or adolescence, since the items used for measurement at such widely disparate stages of development of necessity differ so greatly in kind that their psychological equivalence becomes uncertain. Needless to say, the practice sometimes followed of combining scores or IQ's derived from different tests within a single distribution for the purpose of group statistics greatly increases the chances of errors of this type. This raises the question of the relation between initial scores and increment, as well as the question of the proportion of the final true score that is represented by the common factor of the initial score in determining the correlation between tests.
5. Errors of interpretation resulting from failure to take account of the principle of regression, particularly in its effects on measurements of gain or loss.

These are not trivial matters. The scientific validity of any study is a direct function of the soundness of its experimental techniques, by which it must stand or fall. In Part I, Chapter XI of this Yearbook, the evidence presented in various experimental articles dealing with the effect of environmental stimulation upon the intelligence of children has therefore been subjected to critical review in the hope that such an appraisal may lead to a better understanding of the relations involved. Even though the survey may not lead to the formulation of any very definite conclusions, it should at least help to indicate the nature of the difficulties encountered in studies of this kind and show where further investigation is most urgently needed. In this field, particularly, the formation of hasty judgments from inadequate evidence is perilous. It is well to bear in mind the time-honored adage: "The perception of ignorance is the first step toward knowledge."

CHAPTER XIII

THE PREDICTION OF TERMINAL INTELLIGENCE FROM INFANT AND PRESCHOOL TESTS[1]

John E. Anderson
Director, Institute of Child Welfare
University of Minnesota
Minneapolis, Minnesota

Wellman (19, 20, 21, 22) and Skeels (12, 13, 14, 15) have called some of the assumptions of modern intelligence-test theory into question. Their results, based upon infant or preschool tests as origins, are interpreted to show that a favorable environment produces great changes upward, and an unfavorable environment great changes downward, in *true* intelligence. This contribution to the Yearbook is not concerned with the results of these studies in detail, but is concerned with the reliance that is to be placed upon infant and preschool tests as measures of what is later known as intelligence, *even* if we accept the published findings as correct.

Because of the use of the criterion of age progression in the standardization of infant and preschool tests and the neglect of the criteria of ratings, of measures of scholastic achievement, and of internal consistency that are used with older age levels, infant scales consist very largely of motor items. Infant-scale scores show zero or very low positive correlations with intelligence scores at later ages (2, 6). Recently Richards and Nelson (8, 9, 10), and others, have found that the items in infant scales correlate in different degrees with intelligence-test scores at later ages and suggest that by item analysis and weighting through partial correlation, the correlations of infant tests as a whole with later measures may be raised. In this chapter, I propose the use of test performance at later age levels as a criterion and am concerned with the implications involved in its use. *The proposal as-*

[1] Assistance in the preparation of these materials was furnished by the personnel of Works Progress Administration Official Project No. 665–71–3–69. The more detailed presentation, of which this chapter is a condensation, is: "The limitations of infant and preschool tests in the measurement of intelligence." *Jour. of Psychol.*, 8: 1939, 351–379.

sumes that, by making the best possible prediction of terminal status, we shall also make the best measurement of present status, in so far as our concern is with potentiality rather than with immediate achievement.

The use of such a criterion should result in scales for the earlier years that are more homogeneous in terms of the underlying psychological functions measured and in a clearer delimitation of what is defined as 'intelligence.' It would involve a thorough reworking of our present infant and preschool tests in terms of longitudinal, rather than of cross-sectional, data.

I. Correlations of Successive Measures with Initial and Terminal Status

A number of investigators have commented on the decrease in the correlations of successive measures with an initial measure. Honzik (7) analyzed the relation between mental-test constancy and the interval between tests by means of an age ratio obtained by dividing chronological age on the first test by chronological age on the second, and then correlating this ratio with the correlation coefficients obtained between the tests. The correlation between 22 age ratios and the corresponding r's for *Cal* 1 is $+.92 \pm .02$, and for *Cal* 11, $.78 \pm .06$. Thus higher correlations are found for tests closer together in time and lower correlations for those separated by longer intervals.

In the Bayley (1) study, the correlations of mental tests at successive periods from 1 to 36 months show a striking decrease (from .57 to —.09) as we move away from initial status, and a striking increase (from —.09 to .80) as we move toward terminal status. In the Honzik (7) study the correlations at successive periods from 1.90 months to 7 years decrease from .68 to .42 away from initial status and increase from .42 to .81 toward final status. Table I presents these results.

From the data of the Harvard Growth Study (3), 135 boys and 130 girls were selected, on whom ten-year records were complete and the correlation coefficients were calculated for mental age at each year level, with mental age at 7 years as initial status and mental age at 16 years as terminal status. Unfortunately, the children in this study were not given the same mental tests year after year. This reduces the intercorrelations and makes the data more unsatisfactory than would be data obtained throughout from the same scales. In spite of this deficiency, a trend is clear that justifies further analysis. These results are presented in Table II. For the boys the correlations with initial status decrease from .735 to .582; for the girls, from .651 to .542. For the boys, correlations with terminal status increase from .582 to .901; for the girls, from .542 to .906.

TABLE I. — DATA FROM THE RESULTS OF BAYLEY AND OF HONZIK

Bayley			*Honzik*		
Age in Months	*Initial Status 1 to 3 Months*	*Final Status 27 to 36 Months*	*Age in Years*	*Initial Status 1.90 Years*	*Final Status 7.0 Years*
1 to 3	—	−.09	1.90	—	.42
4 to 6	.57	.10	2.0	.68	.46
7 to 9	.42	.22	2.5	.59	.38
10 to 12	.28	.45	3.0	.47	.56
13 to 15	.10	.54	3.5	.50	.63
18 to 24	−.04	.80	4.0	.46	.66
27 to 36	−.09	—	5.0	.32	.73
			6.0 *	.30	.81
			7.0 *	.42	—

* Stanford-Binet.

TABLE II. — DATA CALCULATED FROM HARVARD GROWTH STUDY

130 Girls			*135 Boys*		
Mean Chronological Age in Years	*Correlation with Initial Status 7 Years*	*Correlation with Terminal Status 16 Years*	*Mean Chronological Age in Years*	*Correlation with Initial Status 7 Years*	*Correlation with Terminal Status 16 Years*
7.41	—	.542	7.44	—	.582
8.40	.651	.584	8.43	.735	.641
9.40	.604	.533	9.42	.697	.581
10.40	.719	.700	10.43	.726	.744
11.40	.668	.728	11.42	.670	.752
12.39	.655	.776	12.42	.642	.790
13.39	.642	.812	13.41	.659	.778
14.39	.632	.822	14.41	.653	.829
15.39	.569	.906	15.42	.606	.901
16.39	.542	—	16.42	.582	—

In the following tabulation the correlations from the Honzik data and the Harvard Growth Study are compared over a six-year span.

With Initial Status						*With Terminal Status*					
Honzik With 1.9 Years		*Harvard With 7 Years*				*Honzik With 7 Years*		*Harvard With 16 Years*			
Age		*Age*	*Boys*	*Girls*		*Age*		*Age*	*Boys*	*Girls*	
2	.68	8	.735	.651		2	.46	11	.752	.728	
3	.47	9	.697	.604		3	.56	12	.790	.776	
4	.46	10	.726	.719		4	.66	13	.778	.812	
5	.32	11	.670	.668		5	.73	14	.829	.822	
6	.30	12	.642	.655		6	.81	15	.901	.906	

The correlations for initial status are with 1.9 and 7 years, respectively; for terminal status, with 7 and 16 years, respectively. Although the data are not so perfect or so comparable as one would wish, it is clear that the coefficients obtained later in the developmental course are significantly higher than those obtained earlier and that the correlations with initial status drop much more rapidly in the earlier ages. Correlations with terminal status build up more rapidly in the earlier ages, but do not reach so high a level within a comparable span. This is also apparent for a different age level if one compares the Bayley and Honzik data in Table I.

II. The Concept of 'Overlap'

We deal here with a phenomenon in which the prediction of final status is based upon a larger and larger proportion of that which is included in the total; that is, scores at 10 years include more of that which is present at 16 years than do scores at 3 years. What then is the nature of the relation between earlier and later measures when successive measurements include larger and larger parts of that which makes up final status and smaller and smaller parts of that which makes up initial status? In order to arrive at the determining principle, the numbers on the faces of playing cards were recorded after shuffling. The cards were again shuffled and the numbers that turned up were added in succession to the results of the first shuffle; the results of the next shuffle were added to the sum of the previous two shuffles, and so on for 16 shuffles. Scores for 96 cases that cumulate from the first to the sixteenth shuffle were obtained. The cumulated scores at each shuffle were then correlated successively with initial score and with final score. A similar procedure was followed using the Tippett [1] table of random numbers to make up a series of 300 scores, cumulated from the first score to the sixteenth. In these the increments at each level are uncorrelated and have similar means and standard deviations. Hence, the cumulated means at successive levels increase by constant amounts. By obtaining the ratio of the means at each successive level to the initial mean and terminal mean, the proportionate amount of overlap can be obtained, and the correlations obtained can be plotted against these ratios.

The formula for handling the problem of overlap is found in the coefficients of determination and non-determination, which measure the amount of association between two measures, or the extent to which the variance in one variable is determined by that in the other variable (4, 6). Since $r^2 + k^2 = 1$, r^2 becomes a coefficient of determination and k^2 a coefficient of non-determination. The percent of overlapping is given by:

$$r^2 = \text{percent overlap or } r = \sqrt{\text{percent overlap}}$$
$$k^2 = \text{percent nonoverlap or } k = \sqrt{\text{percent nonoverlap}}$$

In Fig. I is presented the curve obtained by plotting the formula with the correlation coefficient as the ordinate and the fraction representing the amount of overlap with the initial measure as the abscissa, together with the results obtained from the playing-card and random-number cumulations.

In Fig. II is presented the curve obtained by plotting the correlation of each successive measure with the terminal measure against percent of overlap, together with the data obtained in the playing-card and Tippett number series. With a one-half, or 50 percent, overlap, r is the square root of .50 or .707; with a one-fourth overlap, r is the square root of .25, or .50, and so on. It is clear that the results obtained with the playing-card series and the random numbers fit the formula.

We may then suggest the principle that the earlier in a cumulated series a measurement is taken, the less predictive it will be of final status, and the later it is taken, the more predictive it will be, and we may ask under what conditions it will, or will not, hold. It will hold only if the increments, as in the playing-card series or random-number series, are uncorrelated, or if the increments correlate with previous position in varying, but moderate, degrees. If the increments are perfectly uniform, it will not hold, because the relative positions at suc-

[1] L. H. C. Tippett. *Random Sampling Numbers.* Tracts for Computers. No. 15 (Cambridge University Press: London, England, 1927) 26 pp.

cessive periods will be unchanged, since a constant will be added uniformly to each score. It will not hold if the increments are differential with respect to original scores and always perfectly correlated with them, since only a fanning out of the individual growth curves will

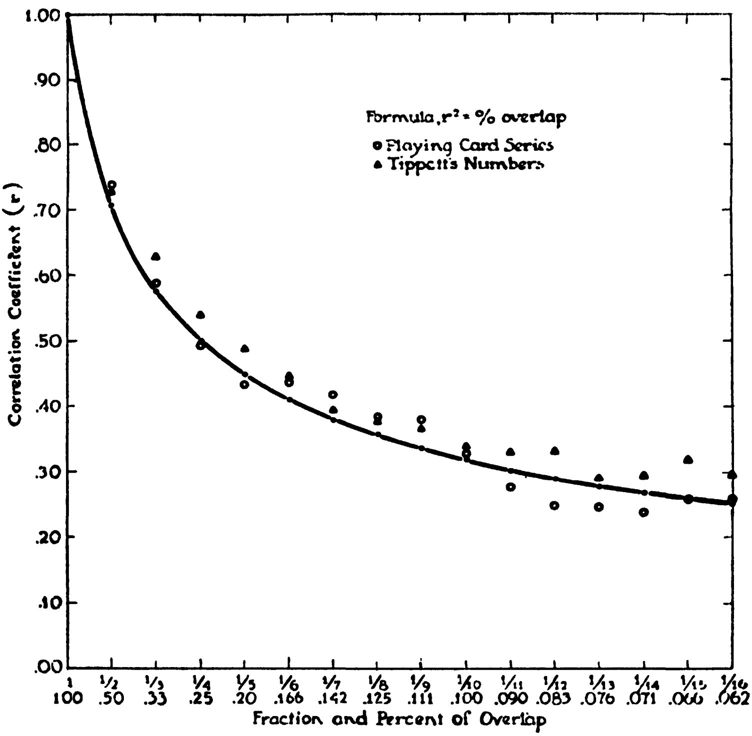

FIGURE I

result. If the correlations with terminal and initial status for mental-test data show the characteristics of the curves obtained from the use of playing cards and random numbers, it would appear that mental growth is either a process in which the increments are not exactly constant or proportional to the original measures, or one in which what is measured at successive periods is composed of different functions that overlap the initial and the terminal measures in content to varying degrees. It is likely that both these alternatives are to some degree characteristic of mental growth as measured by intelligence tests.

Intelligence is a composite of many different functions. We deal not with increments in a single function, but with the resultant or combination of a series of increments spread over a variety of changing and growing functions. Because we lack the means of measuring these functions and their totals in absolute units, we must throw our measurements into an artificial linear framework that uses one year of mental

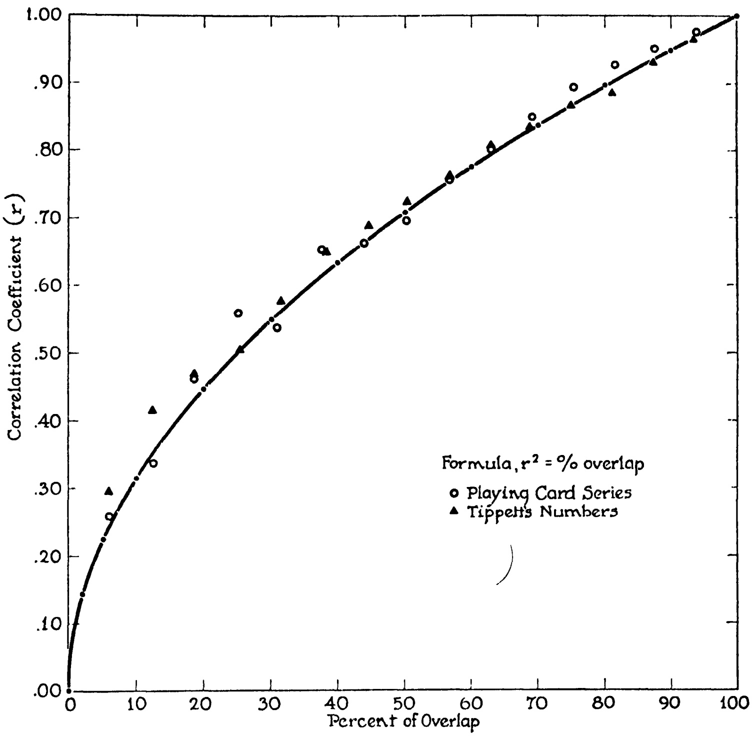

FIGURE II

age as the unit. Underneath this framework there is a progressive differentiation of the functions that go to make up the composite. In a sense, mental age and the intelligence quotient are fictions that gain reality only through their usefulness. In so far as the IQ shows constancy, it measures elements that are common to a number of functions and at any particular level measures only present status. The IQ gains its predictive value from the demonstration that positive and high correlations hold for successive determinations.

Do the correlation coefficients obtained with mental-test results at successive periods in longitudinal studies fit the formula given for determining overlap? Because intelligence-test scores are far from being perfect measures of whatever they measure, this is difficult to answer. Not only are different tests used at different levels in the data available, but also whatever tests are used, the problem of the reliability of the measures arises. The two sets of data analyzed in this study are the Honzik data, obtained in the early years, and the Harvard Growth data at later age levels.

In Fig. III[1] the curves for the decrease in coefficients away from initial status, and in Fig. IV the curves for the increase toward terminal

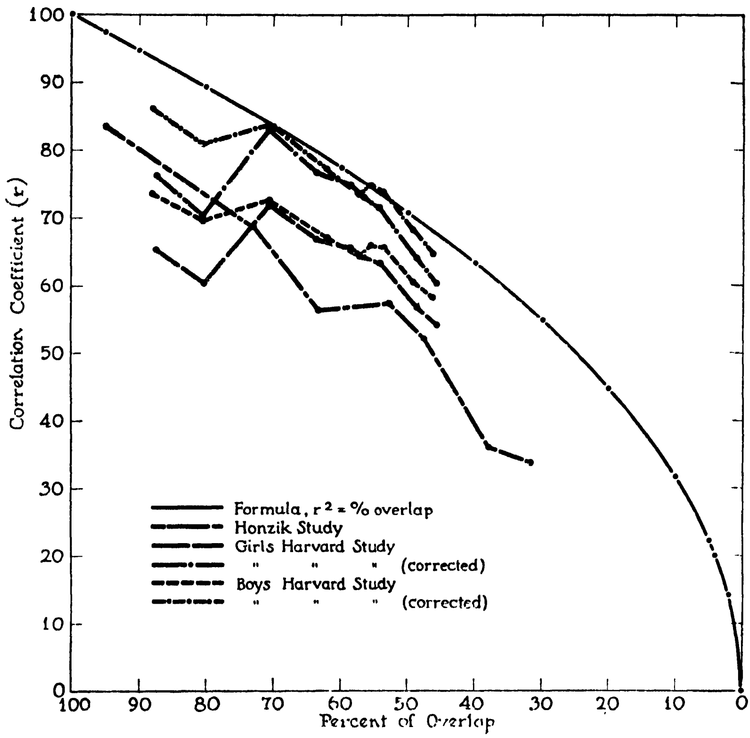

FIGURE III

status, are plotted in terms of the proportions that later are of earlier measures, and that earlier are of later measures, respectively. The Honzik data have been corrected for attenuation, using the reliability

[1] In examining Figs. III and IV, the reader should keep in mind that the initial measurement for the Honzik data is at 1.9 years and the terminal measure at 7 years, whereas the initial measure for the Harvard data is at 7 years and the terminal measure at 16 years.

figures given in her article. For the Harvard Growth data the original coefficients are plotted, together with the same series of coefficients corrected for attenuation, by assuming the reliability at age seven to be .85; age eight, .86; and so on (.01 is added to the reliability with each year of age). This makes the assumed reliability at the terminal measure .94.

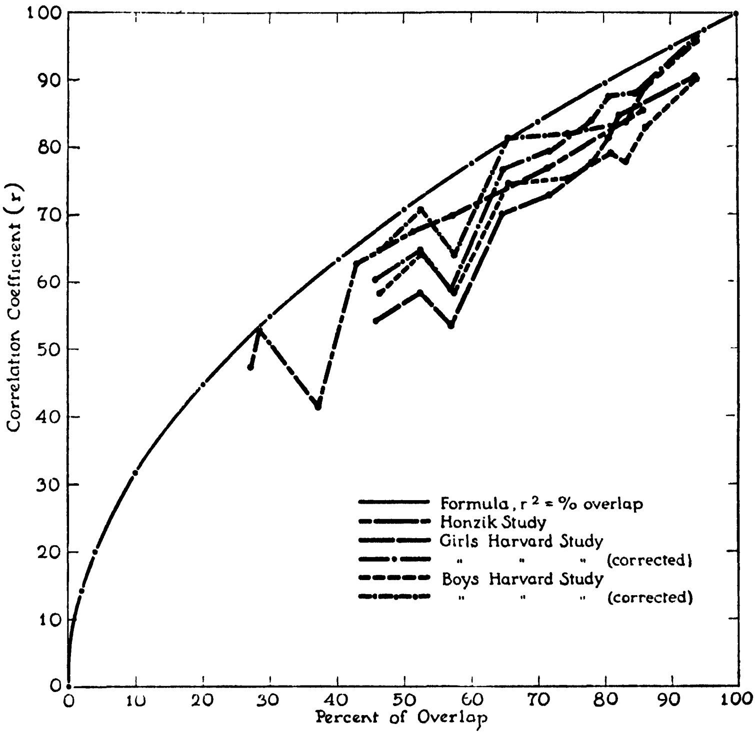

FIGURE IV

It is clear from these figures that the curves obtained approximate that for the formula, but that they lie under it. The Harvard Growth curves are approximately 10 points of r under the curve; the Honzik data, approximately 15 points of r under it. With corrections for attenuation, the Harvard data, from the 9-year level on, come close to the formula.

Whatever questions may be raised with reference to the accuracy with which obtained correlations fit the formula, it is clear that the phenomenon of the increase and decrease of correlation coefficients as we move toward terminal status or away from initial status is one that is related to the amount of overlap. Since the formula gives the relation when increments are uncorrelated, the question why the mental-age curves fall below rather than above the formula curve can well be raised, but at present it cannot be answered. If there were consistent increments from age to age plus commonality of factors in the tests, the coefficients would lie above the formula curve. If the increments were uncorrelated with previous status, but possessed identical means and standard deviations from year to year, the situation in the playing-card and the random-number series would hold and there would be close approximation. Since the curves lie under the expected curve, the factors that are to be sought are (*a*) unreliability of the measures, which we know, when corrected, brings the curves nearer expectancy; (*b*) differences or variations in the content of the tests from year to year, which serve to reduce the intercorrelations; (*c*) small variations in the means and standard deviations of the increments as a result of poor standardization of the tests; and (*d*) negative correlations of increments with previous status.

Since the growing individual does not lose what he already has, the constancy of the IQ is in large measure a matter of the part-whole or overlap relation. The farther he is along in the growth process, the smaller are the increments; thus, with linear age scaling, an increment of one year of mental age at ten years chronological age is one-tenth of what has been attained, while an increment of one year (mental age) at three years is one-third of what has been attained (in total intellectual growth). After ten cumulations in the playing-card and random series, the correlations are .80 or above — figures comparable to those obtained for mental tests, but produced entirely by uncorrelated increments.

Certainly these data suggest the need of a reëxamination of the whole problem of mental age and IQ increments in longitudinal series, using accurately standardized tests and the best possible conditions of testing. In such a reëxamination account must be taken of the overlap factor.

III. Relation to Outside Factors

The relation between successive measures of intelligence and measures of any other functions may change in accordance with the principle of overlap. Successive measures may be viewed as indicators of more or less of that which is associated with the outside variable. A striking example is furnished in the study by Bayley and Jones (2), in which scores on mental tests at successive age levels were correlated with a series of measures of parental and socio-economic status. The following tabulation presents in shortened form several series of the relations reported by them.

Age in Months	*Mother's Education*	*Father's Education*	*Father's Occupation*	*Total S.E. Scale*
4, 5, 6	−.23	−.26	−.26	−.10
10, 11, 12	.06	−.06	.01	.05
18	.12	−.10	−.06	−.10
24	.52	.39	.35	.34
36	.46	.28	.23	.04
48	.50	.37	.31	.22
60	.48	.53	.43	.36
72	.58	.50	.38	.41

The correlations are predominantly zero or negative up to the age of 18 months and then become positive and increase in size until a maximum seems to be reached. The tests seem to measure more and more a factor that is associated with these other variables. How are these results to be interpreted?

Skeels (13), who assumes that his tests are true measures of intelligence, interprets the zero or low correlations between the mental-test scores of children at a median age of 18.8 months and measures of the true mother as indicating an absence of relationship between the child's and the mother's intelligence. In view of the extensive literature on parent-child resemblances, this absence of relationship might well be viewed with suspicion. Likewise the very low coefficients between infants' scores and measures of the foster parents is also suspicious in view of the positive coefficients between .20 and .30 commonly reported as the result of selective placement. If the principle put forth in this paper is correct, it is dangerous to interpret low coefficients of resemblance between parent and child based on infant tests as indicating the lack of parent-child relation in ability. Rather should we ask the question whether the infant tests measure the same functions that tests of older children measure.

It should be pointed out, however, that the changing relations represented in the preceding tabulation could be produced by either environmental or hereditary factors or both, as Bayley and Jones (2) clearly recognize. But one cannot conclude from such changes without further evidence that a training factor only is involved and that a maturation or hereditary factor is not involved.

IV. The Use of an Early Score as an Origin

We may regard *true* intelligence as a parameter to which the test score at a particular time approximates. The problem of the age level at which this approximation gives the most accurate indication of *true* intelligence then arises. But this is a double problem. First, does *true* intelligence change with age, and second, do tests given at particular age levels approximate more accurately or less accurately the parameter intelligence than do tests given at other developmental periods?

To the first of these questions there is no clear answer at present. That there is growth in intelligence cannot be doubted. But is this growth an increase in level; that is, quantitatively, or is it both an increase in level and a change in kind; that is, qualitative as well as quantitative? Thurstone's attempt to determine the age relation of primary mental abilities may make a significant contribution.

To the second question, the answer given by this chapter is that measurement later in the developmental sequence gives a closer approximation to the true potentiality of the individual than does an earlier determination. This is precisely the opposite of the assumption made by Wellman and Skeels, who first match children with respect to intelligence on the basis of early tests and, second, consistently use test results obtained in infancy or early childhood as the origin from which to make calculations of increments and gains. While it may be said that this is the only possible procedure that can be followed in a longitudinal study, nevertheless, *when it is assumed that tests measure the same function throughout a series of longitudinal observations, it is clear that the adequacy with which a particular test measures what it purports to measure has to be determined in terms of its correlation with tests at later ages before it can be used as an origin.* Moreover, because of the extraordinary importance of the first measurement when used as an origin in a series, every precaution should be taken to insure its accuracy and validity and the avoidance of constant errors. We might better regard early tests as approximations, in the same sense as we view performance tests or Army Beta at later ages as approximations and then take every precaution to maximize their correlations with measures of intelligence at a level at which it is known to be well measured.

V. The Problem of Pairing at Early Ages

If two groups of children are paired on the basis of infant or preschool-test performance and the correlations of the scores at successive ages decrease with subsequent test performance, what are the chances that the two groups will be equivalent at subsequent periods? This is an important question for the set-up of a longitudinal experiment. If other factors, such as the hereditary factors, have not been controlled in the sampling, they may operate in successively greater amounts to produce differentiation between the groups; in other words, it is perfectly possible that there may be a spurious matching of the groups and that later they may diverge widely irrespective of the type of training received. Thus, if we suppose two groups matched at the age of 2 years with the coefficients with initial status decreasing by 10 points each year, the amount of overlap may be indicated by the formula for determination as follows:

	Age in Years						
	2	*3*	*4*	*5*	*6*	*7*	*8*
Percent		.90	.80	.70	.60	.50	.40
Overlap		100	81	64	49	36	16

Thus there would seem to be no guarantee that the groups matched at 3 years will still be matched at 6 or 7 years for traits markedly affected by maturation.

A much more effective job of matching can be done from measurements of intelligence at 7 or 8 years, when the correlation coefficients with final status at 16 are of the order of 60 or more, than from measurements prior to 3 years, when the coefficients with a measure of terminal status, such as that at 7 years, are around 40 or below. Further, if correlations with initial status decrease rapidly, it will always be possible to select a number of striking individual cases that will show consistent gains or losses from an early determination as a base, and in extreme cases gains and losses of very large amounts.

It is also likely that the results obtained from the use of matching or control-group techniques in which other variables — such as mother's education or mother's intelligence quotient — are used for selecting the groups will be affected. Thus, if one were to select a group of mothers of low intelligence quotients and a group of mothers of average intelligence quotients, and measure the mental level of their children at successive ages, the phenomenon described in this chapter would result in a series of decreasing intelligence quotients for the children of the mothers with low intelligence quotients, whereas the intelligence quotients of the children from mothers of average intelligence quotients would show little or no change with age. Through regression the mental level of the children of the mothers of low intelligence quotients would fall half-way between the mean mental level of the mothers and that of the general population. But if the original tests measure little of what is finally measured, the early measures of the children from mothers of low mental level would fall about the mean of the population, and then, as intelligence is measured more and more accurately, would move from that position to a true intermediate position. Irrespective of other factors, this trend would appear. This downward trend is very apparent in Figure I of Skodak's study (16) or Figure XII in her other study (17). Likewise, children from an hereditary background of high intelligence would fall near the mean on early tests and show a marked upward trend as intelligence was measured more and more accurately. Perhaps also this phenomenon explains why early training *seems* to be so much more important than does late training in producing the desirable effects found in the Iowa studies.

VI. The Effect of Constant Errors

Throughout this chapter the early tests in a series have been assumed to be reliable. When reliability coefficients were available, it was found that by correcting for attenuation there was closer approximation to the formula for overlap or determination. When, however, a factor is present that produces large errors and these in a constant direction, the correlations with both terminal and initial status or with any outside measure would be seriously affected. Many investigators have raised the question as to the effects of resistance or negativism, which is a much greater problem at the preschool than at older age levels, upon the determination of the IQ of young children. Modern scales eliminate those items and tests that children often refuse to handle and substitute others with more intrinsic interest. The Merrill-Palmer Scale includes a method of correcting for refusals that clearly recognizes the existence of this factor. One of the most interesting investigations of this phenomenon was made by Rust (11), who gave three-year-old children both Kuhlmann-Binet and Merrill-Palmer tests, repeating those tests that were refused (not those failed) on successive days a second, a third, and a fourth time, and more, until the children either definitely passed or failed them. The results from this study as presented in the tabulation that follows indicate significant changes in IQ level, as a result of rescoring on the basis of ultimate success or failure on the items.

On the Kuhlmann-Binet, only 32 of 96 intelligence quotients were unchanged, and in 51 of 96 cases the change upward in the intelligence quotient ranged from 5 to 35 points. Somewhat similar, but not as striking, results were obtained for the Merrill-Palmer when the scores were not corrected for refusals. When corrected for refusals, the changes were slight, as shown by the last column in the tabulation.

Points	*Kuhlmann-Binet*	*Merrill-Palmer* *	*Merrill-Palmer* †
25–35	7	2	0
15–24	18	3	0
5–14	26	38	3
1– 4	14	14	26
Unchanged	31	42	70
Total	96	99	99

* Not corrected for refusals.
† Corrected for refusals.

On the basis of this study negativism seems to result in a constant error in the direction of lowering the intelligence quotient. If such a factor is present, the relation of preschool tests to terminal mental level will be seriously affected. In preparing and standardizing tests for young children, it is both profitable and necessary to subject items to rather rigorous selection in terms of the emotional reactions and resistance they elicit. But however perfectly this is done, there is still a somewhat greater likelihood of variation in rapport for younger children. And it may be pointed out that we must be very careful in interpreting the results of examinations given many years ago when tests were markedly subject to this constant error, *which seems always to be in the direction of a lowered score.* Since this error is most marked in scores at early ages, with the disappearance of negativistic responses with age (a fact clearly established in the literature) a marked gain in IQ would appear.

VII. The Effect of Item Analysis

L. Dewey Anderson [1] has been kind enough to make available some of the results of a study he has made of the relation between early achievement on infant-test items and Stanford-Binet intelligence quotients at five years of age. One hundred children followed in a longitudinal study from 3 months to 5 years were given from 85 to 183

[1] L. D. Anderson. "The predictive value of infancy tests in relation to intelligence at five years." *Child Develop.*, 10: 1939, 203–212.

test items at each of the following ages: 3, 6, 9, 12, 18 and 24 months. The items were selected from the Gesell, Linfert and Hierholzer, and Buhler scales. Dichotomous groups of 15 cases, each differing widely in intelligence quotient at 5 years, were set up and an item analysis made, on the basis of which a selection was made and a new scoring done. The number of items that have significance is low at the early years, but increases at 18 and 24 months. While it was impossible for him to set up an independent group on which to validate the increases in correlation with subsequent status, he did recalculate the coefficients after item analysis and selection both for the total group and for the intermediate group of seventy cases with the cases in the dichotomous groups excluded.

Table III compares the predictive value of a point summation of all the items before analysis and the predictive value after items had been selected to maximize predictive value and reliability. After analysis there is a significant increase in correlation with Stanford-Binet results at 5 years. The results at the early levels are somewhat inconsistent when the dichotomous validation groups are included, but are quite consistent when they are not.

TABLE III. — CORRELATIONS OF INFANT TESTS WITH STANFORD-BINET AT 5 YEARS

Age in Months	*Including Validation Group*		*Eliminating Validation Group*		*Correlation with Status*	
	Before Analysis	*After Analysis*	*Before Analysis*	*After Analysis*	*Initial 3 months*	*Terminal 24 months*
3	.086	.315	.016	.144	—	.341
6	.042	.413	—.021	.148	.592	.303
9	.084	.202	.072	.178	.378	.363
12	.085	.225	.061	.108	.206	.469
18	.200	.365	.126	.179	.241	.484
24	.523	.550	.250	.309	.341	—
N	100	100	70	70		

From these results it is clear that the correlations of infant and preschool tests with subsequent measures of intelligence can be appreciably raised. But, as is shown in the last column of the table, after raising the coefficients the principle of decreasing correlation with initial and increasing correlation with final status still holds.

VIII. Precautions in the Use and Interpretation of Early Tests

From the principles brought out in this chapter several general precautions to be observed in the practical or theoretical use of measurements can be formulated. These are:

1. The earlier the measurements are made, the less reliance can be placed on a single measurement or observation if that measurement or observation is used for predicting subsequent development.
2. The earlier the measurements are made, the greater care should be taken to secure accuracy of observation and record and to follow standardized procedures.
3. The earlier the measurements are made, the greater account should be taken of the possibility of disturbing factors, such as negativism and refusals, that operate as constant errors to reduce scores.
4. Since development is a timed series of relations or sequences, there are for many functions periods below which only a small portion of the function can be measured and above which a progressively larger portion can be measured. Hence, the possibilities of prediction are limited and progression with age is not an infallible indicator of the value of a measurement. Every effort should be expended to secure the most accurate and predictive tests by standardizing tests against multiple criteria rather than against a single criterion, and in particular by the use of later, or terminal, measures as one of the criteria in standardization.

IX. Conclusion

The problem put by this chapter is one that is concerned not so much with the effects of training in the early years upon mental-age scores and intelligence quotients as now measured, as it is with a return of emphasis to the selection of test materials, in order that there will be present in the tests to be devised the highest amount of the stable factor we call 'intelligence,' as distinct from the more unstable factor we call 'achievement.' Modern testing literature recognizes the existence in all measuring devices of factors that do not change greatly with experience — measurement of nature, and of factors that do so change — measurement of nurture. For practical and theoretical purposes, we need both types of measures. The effectiveness of each increases as it becomes a purified measure of what it seeks to measure. If we admit that present measures are inadequate, particularly at the earlier developmental levels, we will facilitate scientific progress and center our attention upon the improvement of our instruments.

Infant tests, as at present constituted, measure very little, if at all, the function that is called 'intelligence' at later ages. Preschool intelligence tests, while they are instruments of some value and usefulness, measure only a portion of that function. Whether it will be possible to develop tests at these levels that will measure more of the function remains to be seen. But it is unfortunate that workers in the preschool period have used age-progression as virtually the only criterion for validating their tests. With a clearer recognition of the importance of other criteria, and particularly of the necessity of maximizing correlations with terminal measures, it may be possible to develop tests of high predictive value and usefulness. But until these methodological problems are clarified, it seems futile to make sweeping generalizations with respect to the nature of intelligence on the basis of present measures in the preschool period, and especially so on the basis of our present measures in infancy.

References

(1) Bayley, N. "Mental growth during the first three years: A developmental study of sixty-one children by repeated tests." *Genet. Psychol. Monog.*, 14: 1933, 1–92.

(2) Bayley, N., and Jones, H. E. "Environmental correlates of mental and motor development: A cumulative study from infancy to six years." *Child Develop.*, 8: 1937, 329–341.

(3) Dearborn, W. F., Rothney, J. W. M., and Shuttleworth, F. K. *Data on the Growth of Public-School Children from the Materials of the Harvard Growth Study.* (Monographs, for Research in Child Development. Vol. III, No. 1: Washington, D. C., 1938) 136 pp.

(4) Ezekiel, Mordecai. *Methods of Correlation Analysis.* (John Wiley: New York, 1930) 427 pp.

(5) Furfey, P. H., and Muehlenbein, J. "The validity of infant intelligence tests." *Ped. Sem. and Jour. Genet. Psychol.*, 40: 1932, 219–223.

(6) Guilford, J. P. *Psychometric Methods.* (McGraw-Hill Book Co.: New York, 1936) 566 pp. (esp. pp. 361–366)

(7) Honzik, M. P. "The constancy of mental test performance during the preschool period." *Ped. Sem. and Jour. Genet. Psychol.*, 52: 1938, 265–302.

(8) Nelson, V. L., and Richards, T. W. "Studies in mental development: I. Performance on Gesell items at six months and its predictive value for performance on mental tests at two and three years." *Ped. Sem. and Jour. Genet. Psychol.*, 52: 1938, 303–325.

(9) Nelson, V. L., and Richards, T. W. "Studies in mental development: III. Performance of twelve months' old children on the Gesell schedule, and its predictive value for mental status at two and three years." *Ped. Sem. and Jour. Genet. Psychol.*, 54: 1939, 181–191.

(10) Richards, T. W., and Nelson, V. L. "Studies in mental development: II. Analysis of abilities tested at the age of six months by the Gesell schedule." *Ped. Sem. and Jour. Genet. Psychol.*, 52: 1938, 327–331.

(11) Rust, M. M. *The Effect of Resistance on Intelligence Scores of Young Children.* (Teachers College, Columbia University, Child Development Monographs, No. 6: New York, 1931) 80 pp.
(12) Skeels, H. M. "Mental development of children in foster homes." *Ped. Sem. and Jour. Genet. Psychol.,* 49: 1936, 91–106.
(13) Skeels, H. M. "Mental development of children in foster homes." *Jour. Consulting Psychol.,* 2, No. 2: 1938, 33–43.
(14) Skeels, H. M., and Fillmore, E. A. "Mental development of children from underprivileged homes." *Ped. Sem. and Jour. Genet. Psychol.,* 50: 1937, 427–439.
(15) Skeels, H. M., Updegraff, R., Wellman, B. L., and Williams, H. M. *A Study in Environmental Stimulation: An Orphanage Preschool Project.* (University of Iowa Studies in Child Welfare, Vol. 15, No. 4: Iowa City, 1938) 191 pp.
(16) Skodak, M. "The mental development of adopted children whose true mothers are feebleminded." *Child Develop.,* 9: 1938, 303–308.
(17) Skodak, M. *Children in Foster Homes: A Study of Mental Development.* (University of Iowa Studies in Child Welfare. Vol. 16, No. 1: Iowa City, 1939) 156 pp.
(18) Thorndike, Robert L. "The effect of the interval between test and retest on the constancy of the IQ." *Jour. Educ. Psychol.,* 24: 1933, 543–549.
(19) Wellman, B. L. "The effect of preschool attendance upon the IQ." *Jour. Exper. Educ.,* 1: 1932, 48–69.
(20) Wellman, B. L. "Some new bases for interpretation of the IQ." *Ped. Sem. and Jour. Genet. Psychol.,* 41: 1932, 116–126.
(21) Wellman, B. L. "Growth in intelligence under differing school environments." *Jour. Exper. Educ.,* 3: 1934, 59–83.
(22) Wellman, B. L. "Mental growth from preschool to college." *Jour. Exper. Educ.,* 6: 1937, 127–138.

CHAPTER XIV

ENVIRONMENT AND THE IQ

GEORGE D. STODDARD
Director, Iowa Child Welfare Research Station
and
BETH L. WELLMAN
Professor of Child Psychology
Iowa Child Welfare Research Station
State University of Iowa
Iowa City, Iowa

In the years since the publication of the *Twenty-Seventh Yearbook* (51) of this Society, on "Nature and Nurture: Their Influence upon Intelligence," the literature on intelligence tests has been voluminous. The present review does not attempt to cover the entire field, but rather to select those studies that appear to make a contribution to the problem of environment in relation to the IQ. With a few exceptions, the discussion is confined to material on this subject published subsequently to the last Yearbook. Occasionally a study published prior to 1928 has been included when, in the judgment of the reviewers, it appeared especially pertinent to the discussion, or when it made a special contribution not duplicated in later studies.

The studies to be reviewed have been grouped under six general headings: (1) home influences, (2) schooling: preschool attendance, (3) schooling: type of school attended, (4) special training programs, (5) general home and school conditions, and (6) institutional residence. A few of the studies could have been included appropriately under more than one heading. In such cases, the reviewers made an arbitrary choice.

I. HOME INFLUENCES

Under this heading are discussed (1) socio-economic status of parents, (2) education of parents, (3) IQ of parents, and (4) children transferred to foster homes. It is recognized, of course, that the child lives in a complex world, subjected to a number of environmental forces simultaneously. He is seldom solely under the influence of the home apart from community influences. This condition, however, more nearly approximates a sole influence at the younger ages than after the child enters school.

It is recognized also that relationships existing between child ability and the categories of home influences just listed must be interpreted from the standpoint of both environmental influences and the inheritance of abilities. Few investigators, if any, hold that the environment has no effect upon intelligence. The question, then, becomes one of determining how great that influence is under specified conditions,[1] and how much variation of influence there is among different environmental conditions.

1. Socio-Economic Status

Comprehensive critical reviews of socio-economic status in relation to IQ have been made recently by Neff (52) and by Lorimer and Osborn (46). The latter review has discussed the subject in terms of its relevance to population problems.

It has been found consistently that groups of children whose fathers are in the upper occupational levels have a higher mean IQ than groups of children whose fathers are in the lower occupational levels. Neff (52) reports that, in general, children of the lowest urban social group (unskilled common laborers) average about 20 points lower than children from the professional classes. After reviewing 63 studies, Neff concludes:

> Rather, we are raising the question as to whether the generally obtained relationship between social status and intelligence test score requires any other explanation than an environmental one. We feel there is strong, perhaps decisive, evidence for a negative answer. In other words, we have tried to show that although individuals at birth may differ in native endowment (a still incompletely settled question), it has definitely *not* been proved that social status of the parent has anything to do with the native endowment of the infant. That a positive relationship later in the life of the individual may develop is hardly denied. But all the summarized studies tend to show that low cultural environment tends to depress IQ approximately to the degree agreed as characteristic of laborers' children, and that a high environment *raises* IQ correspondingly. All, then, of the twenty-point mean difference in IQ found to exist between

[1] To the reviewers, attempts to determine the maximal contribution of environment under optimal conditions appear futile in our present state of ignorance of environmental forces. No one is in a position to state what constitutes an optimal environment for any given individual.

children of the lowest and highest status may be accounted for entirely in environmental terms. (52; pp. 754–755)

A study by Hildreth (26), based on Stanford-Binet examinations of 608 Oklahoma school children in Grades I to VIII, yielded a considerably greater divergence between means of the highest and lowest occupations than the customary 20 points. Hildreth reported the median IQ for children who had been divided into groups according to the following occupational classifications:

Group I. Professional, including doctors, lawyers, teachers, professional engineers.

Group II. Executives and managers, bankers, factory managers, real estate and insurance dealers, store owners.

Group III. Skilled labor and clerical workers, postal clerks, shop and factory clerks, shop foremen, skilled oil-field occupations, carpenters, plumbers, mechanics, glass-blowers.

Group IV. Semiskilled labor, including oil-field pumpers, pipe-line workers, teamsters, truck drivers, farmers.

Group V. Unskilled labor, including day laborers, road menders, hucksters, washerwomen, and those described as " always out of work " or of whom it was said " works at anything he can get."

The median IQ was 97. The median IQ for the different occupational classifications was as follows:

Occupational Group	*Median IQ*
I	113
II	107.5
III	97.5
IV	84
V	75.7

It can be seen that there was a sharp difference in median IQ between adjacent occupational levels, that there was a progressive decline in IQ, that the highest group was superior and the lowest group approached feeblemindedness, and that the difference between the two end groups was 37.3 IQ points.

The most extensive study based on individual examination of child IQ in relation to occupation is that recently reported by Terman and Merrill (76) for composite mean IQ's on Forms L and M of the newly revised Stanford-Binet. Their results are presented for four age groupings: 2 to 5½ years, 6 to 9 years, 10 to 14 years, and 15 to 18 years. Since the pattern was substantially the same at different ages, only the age groups 6 to 9 and 10 to 14 years are used here for illustration. The total number of subjects in all age groups was 2,757.

Occupational Group	*Description*	*Mean IQ* 6 to 9 Years	*Mean IQ* 10 to 14 Years
I.	Professional	114.9	117.5
II.	Semiprofessional and Managerial	107.3	112.2
III.	Clerical, Skilled Trades, and Retail Business	104.9	107.4
IV.	Rural Owners	94.6	92.4
V.	Semiskilled, Minor Clerical, Minor Business	104.6	103.4
VI.	Slightly Skilled	100.0	100.6
VII.	Day Laborers, Urban and Rural	96.0	97.2

Again it will be seen that there was a wide difference in IQ between the professional and the day-laborer groups. The mean differences of 18.9 and 20.3 IQ points are, however, only about half the difference in the Hildreth group. Since children of day laborers are nearly average in intelligence in the Terman and Merrill group, one wonders where, in the population, may be found samplings that would show a low average. And who can claim that *rural owners* carry, genetically, a stronger mental defect than *day laborers, urban and rural?* Obviously, we must look beyond heredity and beyond gross occupational status for the real explanation of these wide differences.

These two studies illustrate the general trend of a relationship between child IQ and socio-economic status, but they also indicate the precariousness of predicting one from the other without supplementary information. They indicate that such prediction is dubious even for group means, to say nothing of individuals.

Jordan (33) found that nearly 48 percent of the children of mill-workers tested in the dull, borderline, and feebleminded classifications. Their mean IQ was 91. The mean IQ for children of farmers was 88, that for children from the professional classes, 105. The total sampling studied was 1,200 school children.

Paterson and Rundquist (56) found that a large percentage of individuals admitted to an institution for the feebleminded, and of applicants likewise, came from the lowest occupational categories. They studied 823 admissions and 516 applicants for admission and compared them with a random sampling of 866 adult males in Minneapolis. The percentages were those tabulated here.

Occupational Category	*Percent in Admis- sions*	*Percent in Random Sample*
I	0.4	3.0
II	4.1	8.0
III	8.1	30.0
IV	31.5	27.0
V	17.0	27.0
VI	38.9	5.0

They also found that the idiot and the imbecile levels were distributed comparably to the random sample, but that higher-grade feebleminded persons were contributed largely by the lower occupational levels. The percentages of admissions in the different IQ classifications that came from occupational Group VI were as follows:

IQ	*Percent*
1 to 20	17.3
21 to 40	20.7
41 to 60	45.3
61 and above	65.2

They believed "that low-grade feeblemindedness is caused, for the most part, by accidental, pathological, non-hereditary factors which would be distributed more or less at random among the various classes of society whereas simple feeblemindedness is transmitted by biological heredity" (p. 123). Another interpretation is possible, however; namely, that a certain percentage of high-grade feeblemindedness is the direct result of inadequate environment, a condition more likely to obtain in the lower occupational levels than in the higher levels.

Bradway (7) attempted to arrive at an explanation for the inverse relationship between degree of feeblemindedness and occupational status of the father. Children at the Vineland Training School were classified according to diagnosis of "primary" and "secondary" etiology. The distribution of paternal occupations for the "primary" cases was piled up toward the lowest occupational class, while that for the "secondary" cases approximated a symmetrical curve. "Primary" cases were those individuals for whom there was no evidence of post-natal causation and whose family background plausibly indicated hereditary transmission of deficiency. It seems probable that occupational level of parent may have been one of the factors taken into account in determining the classification.

Even within a given sampling of population, prediction of a child's IQ from paternal occupation is extremely uncertain, for there is a wide range of IQ's in each occupational classification and much overlapping between classifications. This is illustrated in Hildreth's study by the range of IQ's in each classification, the percentage of children falling above the median of the total group, and the percentage of children falling below average (90 IQ or below):

Occupational Group	*Range of IQ*	*Percent Above Total Median*	*Percent Below Average*
I	70 to 150	84	7.7
II	45 to 150	75	13.6
III	60 to 140	52	31.1
IV	50 to 115	22	69.2
V	45 to 115	7	88.9

In each classification there was a range from below average to superior. Even in the highest group, nearly 8 percent were below average, and 16 percent fell below 97 IQ. Since the means were closer together in the Terman and Merrill material, prediction there based on any occupational group would be even more unreliable.

Byrns and Henmon (9), in a study of 100,820 Wisconsin high-school seniors given the Ohio State University Psychological Test, American Council Psychological Examination, or the Henmon-Nelson Test of Mental Ability, also report wide overlapping between groups. Differences within every group were greater than the differences between groups. More than one-half of the children above the general median were from the laborer and farmer groups, the two which ranked lowest. This was due in part to the larger number of pupils from these two groups, even for high-school seniors who represent a selected group.

Beckham (5) found less spread than the usual 20 points between the means for 1,100 Negro adolescents from three metropolitan areas. However, his subjects may have been handicapped by the ceiling of the Stanford-Binet, since they ranged from 12 to 16 years of age. Sixty percent of the lowest 2 percent of IQ's and 35 percent of the highest 2 percent of IQ's were in the two lowest occupational classes.

A progressive drop in IQ with increasing chronological age by children from the lowest occupational levels has been reported by Skeels and Fillmore (68), Lichtenstein and Brown (44), and Speer

(73). Skeels and Fillmore studied 407 children from underprivileged homes admitted to an orphanage. Children admitted at ages 4 to 8 years had mean Stanford-Binet IQ's in the 90's (with the exception of the five-year-olds, who tested 89.7); from ages 9 to 14 years, the mean IQ's were in the 80's. Siblings on entrance were found to vary in IQ in proportion to the length of time they had remained in their own homes. The differences are shown here in tabular form.

Difference in Length of Time in Home, Years	*Advantage in IQ of Younger Sibling*
1	2.7
2	3.8
3	7.3
4	6.2
5	7.5
6	8.7
7	8.3
8	11.8
9	14.8

Lichtenstein and Brown (44) found a distinct downward trend in the mean IQ's of successive age groups of children in delinquency areas. The average IQ decreased from 99.6 for the 9-year-old group to 79.3 for the 15-year-old group. The children were in Grades IV to IX.

Speer's (73) study is reported in full in Part II of this Yearbook. The children of feebleminded mothers who were removed from their homes before the age of 3 years had a median IQ of 100.5, while those remaining with the mother until 12 to 15 years had a median IQ of 53.1.

It is difficult to reconcile these findings with Terman and Merrill's (76) report that older children of the lowest occupations are of about average intelligence. The studies reported above dealt with a selection downward within the lowest occupations, but it is probable that they are representative of a group large enough in number to exert an appreciable effect upon the mean for the total group. To what extent the differences between Terman and Merrill and the other studies are due to the fact that the studies reporting decreases used the 1916 revision is not clear. It is possible that some correction should be made for defective standardization of the 1916 revision at the older ages. However, it is not clear how much the correction should be. Studies reporting on the IQ in relation to birth order have not generally given evidence of a lower IQ for older siblings.

If a correction is applied to IQ's utilized in occupational studies, it should be applied to *all* studies using children at older ages. Studies that have reported constant IQ's would, with such a correction, show increases in IQ, while studies reporting gains would show even greater gains.

From the hereditary standpoint, it might be argued that children from the lowest occupational levels represent a type of heredity in which progressive deterioration is to be expected. If this interpretation is accepted, it precludes a theory of constancy of the IQ. It follows that a classification such as the following would be necessary: children with Heredity Type A would decrease in IQ, those with Heredity Type B would not change, and those with Heredity Type C would increase. Such a theory, so far as the reviewers know, has not been seriously advanced. The more plausible interpretation of the decreases is that, as the child grows older, the environment is progressively ill suited to adequate mental growth.

During infancy there appears to be little or no relationship between mental-test ability and the occupation of parents. This lack of relationship has been shown by Furfey (20), Bayley and Jones (4), Skodak (71), and Dubnoff (16). Furfey and Muehlenbein (21) later reported a higher relationship at 5 years of age between Stanford-Binet and socio-economic status than between Stanford-Binet and the Linfert-Hierholzer infant tests. Bayley and Jones (4), who tested the same group at successive ages, found some tendency for the correlation to increase with age, although the trend was not completely consistent. Goodenough (23) found differences in mean IQ on the Kuhlmann-Binet for children 2 to 4 years of age who were classified according to father's occupation.

Differences in mean IQ of children on entrance to a nursery school that are related to occupation of father have been reported by Coffey and Wellman (13) and Bird (6). Coffey and Wellman showed that occupation of father did not, however, account for the gains made by the children while attending the preschool.

Kawin (34) found a difference of 18.0 IQ points on the Stanford-Binet between a group of preschool-age children of high and a group of low socio-economic status. The difference in Merrill-Palmer IQ was 5.7 points. Wellman (89) found no difference in Merrill-Palmer IQ for children attending the preschool laboratories whose fathers were in occupational Class I and those whose fathers were in Classes

II to IV. The same children showed a difference of 6.6 points in Binet IQ.

Skodak (71) studied the IQ of adopted children of preschool age in relation to the occupation of foster father. Since selective placement occurred to some extent, the crucial comparison was one in which true-parent occupation was held constant and foster-parent occupation varied. Children, all of whose true fathers were in the seventh (lowest) occupational classification, were compared. Those whose foster parents were in the three upper occupations were higher in IQ than those whose foster parents were in the lower four occupations.

Lorimer and Osborn (46) have presented some interesting analyses of the Freeman (19), Burks (8), and Lawrence (42) data, which tend to show that the IQ of foster children varies with both true-parent and foster-parent occupation. In the main group of the Lawrence study there was apparently non-selective placement, as the children were placed in boarding homes in rotation. No child was older than one year when placed. The following mean Stanford-Binet IQ's were obtained by children whose true parents (midparents) were classified according to occupations A and B (tradesmen, clerks, highly skilled artisans, and a few professional workers and farmers), C (skilled and semiskilled), and D and E (unskilled or irregular workers):

Occupation	*Number*	*IQ*
A and B	53	103.5
C	107	100.9
D and E	67	96.2

The difference in IQ between the highest and lowest classifications was statistically significant. Similar differences were obtained on cases given the Simplex test.

Lorimer and Osborn's analysis of Freeman's children who were placed under 2 years of age revealed distinct differences between legitimate and illegitimate children according to true parentage. The differences between the two groups were not shown in terms of occupation, but in percentage of mothers who were "mentally defective" and percentage who completed a grammar school. In the legitimate group 51 percent of the mothers were considered "mentally defective" and 30 percent had completed a grammar school. In the illegitimate group the percentages were, respectively, 25 and 60. The mean IQ's of the children were as follows:

	Legitimate		*Illegitimate*	
Grade of Foster Home	*Number*	*IQ*	*Number*	*IQ*
Good (score 23–30)	10	106.5	45	112.4
Average (score 16–22)	19	98.3	39	104.9
Poor (score 8–15)	16	83.4	27	95.6

Another point brought out by Lorimer and Osborn was that 36 own children living in the same homes had a mean IQ 9.6 points above the illegitimates.

Also, Burks' (8) group of foster children, all of whom were placed under one year of age, had a mean IQ of 107.4, whereas the own children had a mean IQ of 115.4.

Present evidence, then, appears to indicate that both true parentage and foster home are important factors in the child's mental development. It is to be noted that children placed in adoptive homes tend to be superior in IQ. In Leahy's (43) group the mean was 110; in Skeels' (66), 115; and in Skodak's (71), 116; in Hildreth's (27) group the median was 103. In none of these studies was anything apparent in the case histories on true-family background to justify the expectancy of such superiority. When children are placed in boarding homes, the IQ's are somewhat lower, as would be expected if environmental differences are important. (See Snygg, 72, and Speer, 73.)

2. Education of Parents

Of all the criteria utilized in parent-child relationships, the education of parents is perhaps the most elusive. If we did not have compulsory schooling up to 14 or 16 years, the criterion would have more meaning; then presumably only the able would stay in school, and the others would drop out. As it is, almost everybody stays in school; in many states, it is school or a state institution for every child up to age sixteen. Similarly, for grades completed: in rural and village schools especially, it is common practice to 'promote' children. Consequently, they usually reach the upper elementary grades, regardless of mental disabilities.

The relationship between education of parents and IQ of child has usually been expressed in terms of correlations. These are often difficult to interpret. Goodenough's (22) correlations for separate age groups 2, 3, and 4 years tended to yield higher relationships for both boys and girls at age 3 than at 2 or 4. There was no clear difference between correlations with mother's education and father's education.

The range of correlations was from .09 to .69. Van Alstyne (82) reported a correlation for 75 3-year-old children of .60 ± .05 between child's IQ and mother's education and of .51 ± .06 with father's education. Bayley and Jones (4) showed a lack of relationship until after 18 months of age. At 24 months and after, up to 72 months, the correlations ranged from .40 to .58 with mother's education, from .21 to .53 with father's education, and from .38 to .59 with midparent educacation. The correlations with mother's education tended to be higher throughout than those with father's education. The authors pointed out that this would be expected on a nurtural basis because of greater association of child with mother. They also suggested that it might be a chance characteristic of the small sample. Their final conclusion was that:

> the increasing correspondence between mental score and environmental variables is not necessarily attributable to the influence of the environment; it may equally well be a phenomenon of infant development, that inherited parent-child resemblances become evident only after a certain stage in the process of maturation has been reached. Evidence can be adduced in favor of each of these interpretations; the probability is that each has some validity, and that the growth of children involves both an increasing assimilation of environmental pressures and an increasing manifestation of complex hereditary potentialities. (p. 339)

3. IQ of Parents

Van Alstyne (82) reported a correlation of .54 ± .06 between child IQ at 3 years of age and mother's IQ. This is about the magnitude of the usual relationship reported for school-age children.

Little relationship between adopted-child IQ and true-parent IQ was reported by Skeels (66), Snygg (72), Skodak (71), and Speer (73). Practically a zero correlation (.04) was found by Skeels (66) between IQ of mother and child IQ at the preschool ages for 78 children placed in adoptive homes under 6 months of age. The mean IQ of 22 children whose true mothers tested below 80 IQ was 116.5.

Snygg (72), working independently at Toronto, came to the conclusion that, in a group removed from their homes in infancy, there was little possibility of predicting the IQ of any child from the IQ of the mother, although mothers testing under 70 did have the largest percentage of subnormal children. The percentages of children testing under 70 were, respectively, for IQ classifications of mothers: under 70, 12.2 percent; 70 to 79, 2.6 percent; 80 to 89, 3.2 percent; and over 90, 4.0 percent. The mean IQ of the children was 95.2, of the mothers 78.3. The IQ's of the children were said to be representative of all children tested by the clinic. The children were removed from their homes at an early age and placed in paid boarding homes whose keepers did not rank high in occupation or income. Correlations between child IQ and mother's IQ approached zero at every age and did not increase with chronological age up to 5 years. There was little evidence of the later emergence of an inherited child-parent resemblance referred to by Bayley and Jones. The correlation for 70 children 5 years of age or older was .12.

Skodak (70) made a special study of 16 children whose true mothers were feebleminded. These children had been removed from their mothers and placed in adoptive homes before 6 months of age. The range of IQ's of the mothers was 54 to 74. The children were above average in intelligence. On examination at a mean age of 2 years, 6 months, their mean IQ was 116.4; on retest at 5 years, their mean IQ was 107.6. The 8 children placed in foster homes in the upper 3 occupations had a mean IQ on retest of 111.4, while the remaining 8 children placed in foster homes in the 4 lower occupations had a mean IQ of 103.9. No child was borderline or below.

Speer (73) has reported in Part II of this Yearbook a similar type of study of a larger number of children whose mothers were feebleminded. The children were removed from their mothers, but were not placed in adoptive homes. Their pattern of mental growth was definitely superior to that of the group who remained with their own mothers. Both Skodak and Speer concluded that there is no reason why physically normal children of feebleminded mothers may not be placed for adoption at an early age.

Likewise, Wells and Arthur (91) concluded that the child of feebleminded parents placed in a foster home had an advantage over the child who was left to grow up under the care of the feebleminded parents. Their groups consisted of somewhat older children, one or both of whose parents tested 75 IQ or less. One group of 100 children remained with their parents. Another group of 100 children was placed in licensed foster homes but was not placed for adoption. The own group had a mean IQ of 81.1 at 6 years, 7 months, and 74.4 on retest at 12 years. The foster group had a mean IQ of 79.7 at 5 years, 6 months, and 81.1 at 10 years, 5 months.

Correlations between IQ of foster children and IQ of foster parents are usually lower than for own children living with their parents.

Leahy (43) reported a correlation of .51 between own child IQ and parent's Otis score, while the corresponding correlation between adopted child and foster father was .15 and foster mother .20. The age range in each group was 5 to 14 years. The adopted children were placed in their adoptive homes before 6 months of age. Own children in their own homes were matched with the foster children on age and sex. Foster parents and the parents of the own group were matched on occupation, school attainment, race, and residence in communities of 1,000 or more.

Leahy (43) interpreted these and additional correlations between child IQ and parent and home measures as evidence that heredity is more important than environment in determining child IQ. Speaking of the own group, she says:

> In the case of the latter group heredity and environment are both operative. Hence variance in intelligence is accounted for by variance in heredity and environment combined to the extent of about 25 percent (square of r .50). In the adopted group, however, where environment is functioning independently of heredity, variance in intelligence is accounted for by variance in environment only to the extent of about 4 percent (square of r .20). If we neglect whatever artificial heredity selective placement of adoptive children may have introduced into the data, these coefficients are clear evidence of maximum variance in intelligence with variance in environment. (p. 284)

There is, however, serious question as to the feasibility of attempting to separate the influence of heredity and environment by any means. There is the additional question as to the adequacy of a resemblance technique (correlations) for revealing changes in intelligence brought about by environment.

Specific objections to the procedures of the nature-nurture studies are given more fully in Chapter II on "The Meaning of Environment." Briefly, they cluster about the following:

1. Change in IQ cannot be inferred accurately from end correlations.
2. Measures of heredity are inadequate.
3. Measures of environment are inadequate.
4. Environmental factors outside the home are neglected.
5. The dynamic environments provided by foster and own parents may differ.
6. Maximal effects of environment cannot be determined unless environment is given a maximal chance (a condition not known to obtain in any study).

4. Children Transferred to Foster Homes

Children tested before and after placement in foster homes gained in IQ in the studies of Skodak (71), Schott (64), and Wells and Arthur (91). Skodak (71) found that gains were not made while the children were held in an orphanage prior to placement, but that there were subsequent gains after placement in adoptive homes. Sixty-five children received a placement test at a mean age of 3 years, 4 months. Their mean IQ was 98.5, with a range from 70 to 156. On retest after at least one year of residence in the foster homes, their mean IQ was 104.2, with a range from 80 to 165. The change between two preplacement tests for 20 of the children was 1.3 points.

Schott (64) reported a median gain of 5.8 points in IQ for 74 children in foster (boarding and adoptive) homes. The reason for retest was in many cases failure to adjust in the boarding homes. The interval between tests ranged from 1 month to 4½ years, with a median of 13 months. The age range on first test was from 18 months to 12 years; the median age was 3 years. Schott reported that the gain was not statistically significant and concluded, "Our data indicate little hope of raising the IQ appreciably by foster home placement." Thorndike (77) in a note concerning the Schott study, pointed out an error in calculation, which changed the gain to a significant one. In any case, children failing to adjust may not show the same tendency to gain as children making satisfactory adjustment.

Wells and Arthur (91) found a loss in IQ of 6.7 points by 100 children of feebleminded parents who remained at home, and a gain of 1.4 points by 100 children, also of feebleminded parents, who were placed in foster homes (not for adoption). The age of own children at first test was 6 years, 7 months, and at retest 12 years. The age of foster children at first test was 5 years, 6 months, and at retest 10 years, 5 months. The mean IQ's on first test were 81.1 and 79.7. The higher the IQ of own children at first test, the greater was the mean loss in IQ. The changes were as follows:

	Median Change	
IQ	*Own*	*Foster*
Under 75	0	+11.0
75 to 84	− 4.3	+ 3.5
85 to 94	−11.5	− 7.0
95 or over	−16.7	−14.3
Total	− 6.7	+ 1.4

Children placed under 5 years of age made a gain of 8.2 points; those placed after that age lost 4.5 points.

II. Schooling: Preschool Attendance

Studies of the effects of nursery-school attendance upon intelligence have been made by Barrett and Koch (3), Kawin and Hoefer (35), Waring (84), and Wellman (85, 86, 87, 88, 89). Barrett and Koch (3) studied two orphanage groups matched as to sex, age, initial score on Merrill-Palmer test, and length of institutional experience. After an interval of 6 to 9 months, in which one group of 17 children attended nursery school, retests were made. The mean initial IQ's were 91.7 and 92.6. The preschool group gained 20.9 points in IQ; the control group, 5.1 points. It appeared that the nursery-school experiences had resulted in accelerating the IQ's of these children.

Kawin and Hoefer (35) matched 22 pairs of children applying for admission to two nursery schools. The children who were admitted were matched with non-admitted children in age, mental age on the Merrill-Palmer scale, and sex. After 6.8 months, the nursery-school group was still practically identical with the control group. The nursery-school group had gained 11.4 months in mental age; the control group, 11.5 months. Percentile rank changes were 24.6 and 26.0, respectively. The nursery-school group thus did not show greater growth associated with their nursery-school experience.

Waring's study (84), while based upon a larger number of children, was reported very briefly. The children were in attendance at the Cornell University Nursery School. As much gain in mental age on the Merrill-Palmer scale was made during 5 months of preschool attendance by 103 children as was made in 8 months of vacation time by 21 of the children. A gain of 10 months in mental age on the Stanford-Binet was made by 83 children in 5 months of preschool attendance, as compared with a gain of 12 months in mental age by 52 children in a 9-month period that included vacation. The gain in Stanford-Binet IQ from fall to spring was 8 points.

Wellman (85, 86, 87, 88, 89) reported results of retests of preschool and non-preschool children in a series of published studies, the results and major conclusions of which are summarized in Part II, Chapter XXVI, of this Yearbook. These studies utilized results on a large number of children attending the preschool laboratories of the Iowa Child Welfare Research Station. In addition, in the same chapter is summarized the work of three students: Coffey (12), Kounin (38), and

McCandless (50). Gains in IQ were found to be associated with periods of preschool attendance. The gains were found to persist throughout the later childhood ages. The kind of life these children experienced is described in detail by Updegraff and others (81) in *Practice in Preschool Education.*

In Part II are published also several investigations dealing with the effects of nursery-school attendance upon the IQ. Of these, the studies by Bird (6), Goodenough and Maurer (24), Lamson (41), and Olson and Hughes (54) concluded that nursery-school attendance did not accelerate the IQ. Changes in IQ associated with nursery-school attendance are reported by Starkweather and Roberts (75). A small, statistically non-significant difference between nursery-school and non-nursery-school children is reported by Frandsen and Barlow (18). Reymert and Hinton (60) found gains by children at the nursery-school levels, although they did not find gains at the later ages.

The main conclusions of Goodenough and Maurer (24) are based upon IQ's derived from application of the Minnesota Preschool Scale. Almost no data on this scale have been reported in the literature. In their study, the Stanford-Binet IQ at 5½ years is considerably higher than the initial Minnesota IQ. It should be noted, too, that Goodenough and Maurer's explanation of the increasingly higher IQ's of the non-nursery-school group on the basis of selection rests on the theoretical assumption that parent-child resemblances increase with age.

III. Schooling: Type of School Attended

Two studies comparing the mental growth of children attending different elementary schools and high schools have been published by Wellman (87, 88). The results and major conclusions of these are summarized in Part II, Chapter XXVI.

Lamson (40) did not find a change in mean IQ in children who participated in "a rich and vital school curriculum." One hundred forty-one children who had been members of the fourth grade were divided into groups that were studied over periods of 2, 3, and 4 years.

In the Newman, Freeman, and Holzinger study (53) comparisons were made among 19 pairs of identical twins who had been reared separately over a long period of time. All these twins had been separated in early childhood at chronological ages from 2 weeks to 6 years, but chiefly in the first year. They had been separated up to the time of their examination, although in one or two cases they had seen something of each other, or had corresponded.

Environmental ratings on educational, social, and physical influences were prepared on the basis of the estimates of five judges. From the measurements of these twins, the authors present a remarkable series of correlations as follows:

Items Correlated	*Correlation*
Weight with Physical Environment	.60
Binet IQ with Educational Environment	.79
Binet IQ with Social Environment	.51
International Test with Educational Environment	.46
International Test with Social Environment	.53
Stanford Educational Age with Educational Environment	.91

They concluded that " differences in educational and social environment produce undeniable differences in intelligence and scholastic achievement as measured by our tests " (p. 341). According to these data a very substantial amount of all the Binet IQ differences in identical twins reared apart was attributable to educational and social differences. Since these differences were found not to have such effects upon standing height and other physical characteristics, where the correlations ranged from —.26 to .23, one may say that differential environment affects what one would expect it to affect; namely, intelligence and educational achievement, rather than basic bodily proportions. To quote the authors of the study:

> . . . the environmental factor, operating alone in the case of separated identical twins, can produce differences as great as or greater than those produced when both hereditary and environmental factors operate within twin families, as in the case of fraternal twins. (p. 348) [1]

This important study has already been the subject of intense controversy. A detailed criticism by McNemar (48) led to a thorough rebuttal by Holzinger (29) and a brief counter-rebuttal by McNemar (49). Now that the ' sound and fury ' have died down, Holzinger appears justified in concluding that the procedures and findings of the study remain reliable and valid.

In Part II of this Yearbook, R. L. Thorndike (78) reports upon retests of children attending different schools. Rusk (63) suggests in connection with the lack of difference in IQ between urban and rural children in Scotland that " perhaps nowhere has scholastic opportunity been more evenly equated than in Scotland."

[1] Quoted by permission of the University of Chicago Press.

IV. Special Training Programs

Changes in the IQ's of high-grade feebleminded and dull-normal children who have been introduced to special training programs have been reported by Otis (55), Risden (61), and Kephart (36). Otis studied the changes in IQ of feebleminded girls in a state school, whose median IQ was 68 and whose median life age was 24. Twenty-five girls joined a reading club and made written reports of books read. When 5 books were completed, a certificate was issued. Fifteen girls received 5 or more certificates; their gain in IQ was 7.2 points. Ten girls received zero to 4 certificates; their gain in IQ was 5.6 points. Fifteen girls not in the reading club made a gain of 4.8 points. Otis concluded that training in reading and book work affected the results of the Stanford-Binet tests, and that improvement in IQ was possible for some cases after the age of 16 years:

> Many girls are called feebleminded and sent to institutions on account of economic conditions or sex difficulties who do not test high because of lack of training in language, and such cases are improvable. (p. 315)

Risden (61) worked with 20 children who were misfits in school. They ranged in age from 10 to 13 years and in IQ from 64 to 86, with a mean of 77. A summer's program of training was arranged, including detailed observation and analyses of the wants and needs of the children and explicit hypotheses as to the nature of a training program calculated to fit these needs and to stimulate intellectual activity. There was a median gain in Stanford-Binet IQ of 5.5 points, with a range from 0 to 20 points. On the Cornell-Coxe scale the median gain was 17 points with a range from −8 to +57 points.

Kephart (36) worked with a group of 16 boys, residents in a "self-determining" cottage in a training school. At the beginning of the experiment they ranged in age from 15 to 18 years. Their mean initial IQ was 66.3, with a range from 48 to 80. The aims of the special program were to stimulate constructive activity, to require the child to work out for himself a rational solution for social problems as they arose, and to give informal training in solving abstract problems. Emphasis was placed on the child's own original development of means to an end rather than on the finished product. At the end of the experimental period, after an interval of 18 months, the mean IQ was 76.4, a gain of 10.1 points.

The results of Wheeler (92) on children of borderline intelligence attending public schools indicated a tendency to decrease in IQ.

Trainor (79) tested 30 college sophomores on the Detroit Intelligence Test, Advanced Form; then submitted them to six weeks' training in semantic methods and retested them. The experimental group gained 36 points; the control group, 6 points. Trainor states that the gain of the experimental group moved them from the 62nd percentile to the 96th percentile of the national norms.

In Part II of this Yearbook are reported studies by Pritchard, Horan, and Hollingworth (58), and by Skeels (67). Pritchard, Horan, and Hollingworth (58) conclude that their program failed to alter significantly the intellectual pattern of dull-normal children attending special classes. Skeels (67) reports large gains in IQ by a group of young children placed singly or by two's in wards with feebleminded older girls in an institution, and large losses in IQ by another group of young children in an orphanage. McCandless (50) reports better intellectual progress for a small group of very superior preschool children in a special educational program than for preschool children experiencing only the usual preschool program. This study is briefly summarized by Wellman in Part II.

V. General Home and School Conditions

Studies of children in isolated mountain communities show a marked and progressive decrease in IQ with advancing chronological age. Younger children (age 6 and 7 years) are average or dull-normal, while older children (14 to 16 years) tend to be feebleminded. The steepness of the drop appears to depend on the degree of isolation of the community. Sherman and Key (65) reported the following IQ's:

Age in Years	*Pintner-Cunningham*	*National Intelligence*
6 to 8	84	—
8 to 10	70	—
10 to 12	53	66
12 to 14	—	67
14 to 16	—	52

They found different patterns of growth in different communities. To quote:

> Since the ancestry of the children of all the Hollows came from the same stock, the claim cannot be made that some of these mountain people are

'degenerate' and therefore their children are expected to be retarded intellectually, a claim too often advanced for the supposed inferiority of isolated mountain children. Furthermore, as has been shown in this paper, the young children of the various Hollows do not differ greatly in intelligence, whereas great differences are found between the older children of the different Hollows. The only plausible explanation of the increasing difference with increasing age is that children develop only as the environment demands development. (p. 289)

Hirsch (28) studied 1,945 subjects. The median IQ at 5 and 6 years was 86.6. There was a progressive drop with age up to 13 years, where the median was 73.1. At 14 years the median was 74.6. Hirsch states: "The slow decline of I.Q. in the age-groups is for the most part due to environmental factors" (p. 222).

Wheeler (93) reported on 1,147 children in public schools said to be fairly representative of isolated mountain children in Tennessee. The median IQ on the Dearborn test at 6 years of age was 94.7. There was a progressive decrease with age. The median at 16 years was 73.5. Wheeler says:

It seems from the Dearborn tests that when a child enters school he is only a few points below normality, but as he continues on through different years the median intelligence of the mountain child falls into the lower limit of dullness or bordering near the line of demarcation between dullness and feeblemindedness. (p. 359)

Asher (2) used the Myers Mental Measure, National Intelligence Test, and the Herring and the Stanford revisions of the Binet. The median IQ on the Myers Mental Measure at 7 years was 83.5. There was a progressive drop with age up to 15 years, where the median was 60.6. For children 16 and older, the median was 65.0. Asher concluded that intelligence tests standardized on urban children are not adequate for measuring the kind or amount of intelligence of Kentucky mountain children.

Edwards and Jones (17) studied 247 North Georgia mountaineers. The mean IQ at 7 years of age was 108. At 15 and older it was 70, an average drop of 38 IQ points from the 7-year level.

Klineberg (37) [1] investigated selective migration as an explanation of the higher IQ of the northern Negro over the southern Negro. He studied children whose parents migrated to New York from Birming-

[1] The reader will find other comments on Klineberg in Chapter IX and Chapter XI.

ham and Nashville. The school marks obtained by the children before migration were very slightly below those of the general population of Negro school children in these southern cities. He was unable to find any evidence of a mental selection with respect to the children involved. The intelligence quotient of the children increased with increasing length of residence in New York City. He compared the mean intelligence on the National Intelligence Test of 517 twelve-year-old boys classified according to their length of residence in New York City. A similar study was made of 619 twelve-year-old girls. The scores for boys rose from 64 (one to two years' residence) to 87 (twelve years' residence, New York born); the corresponding scores for girls were 71 and 98. Additional studies were made with Stanford-Binets. Klineberg declares:

> There seems to the writer to be no reasonable doubt as to the conclusion of this study. As far as the results go, they show quite definitely that the superiority of the northern over the southern Negroes, and the tendency of northern Negroes to approximate the scores of the Whites, are due to factors in the environment, and not to selective migration.
>
> There is, in fact, no evidence whatever in favor of selective migration. The school records of those who migrated did not demonstrate any superiority over those who remained behind. The intelligence tests showed no superiority of recent arrivals in the North over those of the same age and sex who were still in the southern cities. There is, on the other hand, very definite evidence that an improved environment, whether it be the southern city as contrasted with the neighboring rural districts, or the northern city as contrasted with the South as a whole, raises the test scores considerably; this rise in 'intelligence' is roughly proportionate to length of residence in the more favorable environment. (p. 59) [1]

McAlpin (47) found an increase of 6 or 7 points in IQ made by Negro children migrating from the south into the District of Columbia.

Rural children were found by Jones, Conrad, and Blanchard (31) to be somewhat inferior in IQ. The median IQ for 351 children living in northern New England was 92.3. No general adverse cumulative effect upon performance was discernible. The rural child was handicapped on specific test items.

VI. Institutional Residence

Studies of mental growth of orphanage children have been made by Lithauer and Klineberg (45) and Crissey (14). Lithauer and Kline-

[1] Quoted by permission of the Columbia University Press.

berg (45) found a median gain in IQ of 6.3 points by 120 children examined shortly after admission to an orphanage and retested after a period ranging from 3 to 57 months. The mean IQ on first test was 82.3. There was a greater gain by the younger subjects. Children 6 years of age or younger on admission gained 8.5 points; those 8 years of age or older lost 2.6 points. The authors believed that the change in environment had not been marked and that the 'superior' environment of the orphanage would not rank very high in objective terms.

Crissey (14) studied changes in IQ of children in two state orphanages. Children classified as borderline or moron tended to remain constant or to show slight gains; normal and superior children showed consistent losses. When 16 normal and dull-normal individuals were transferred from an institution for the feebleminded to the orphanages, their changes followed the pattern of changes of other orphanage subjects. Every transferred child with an IQ above 93 decreased and every child with an IQ below 85 gained. The mean change of the higher group was from 100.3 to 94.1; of the lower group, from 81.3 to 88.4. The intervals were 17.7 and 15.0 months.

Skeels, Updegraff, Wellman, and Williams (69) studied two matched groups of orphanage children of preschool age over a three-year period. The experimental group attended the orphanage preschool; the control group did not. Large decreases in IQ were associated with long periods of orphanage residence in the case of members of the control group who tested in the upper half of the group (80 IQ or above). Children of similar levels who attended the preschool did not change substantially. The results are summarized briefly by Wellman in Part II, Chapter XXVI.

In Part II, Chapter XVII, Reymert and Hinton report that little change in IQ was made by children at Mooseheart.

Numerous studies in the literature have indicated a tendency for institutionalized feebleminded to decrease in IQ. Among the more recent studies are those by Chipman (11), Arthur (1), Prouty (59), Hilden (25), Davenport and Minogue (15), and Woodall (94), though in the last-named the yearly decrease for a group of 497 subjects was only 0.45 IQ point. Kuhlmann (39), whose study in 1921 deserves mention because of its comprehensiveness, reported an average yearly decrease of 2.19 points for borderline cases, 1.21 for morons, 1.04 for imbeciles, and .37 for idiots. Hilden (25) found a decrease after one year of 1.8 points, and after six years of 7.8 points. Davenport and Minogue

(15) reported about a two-point yearly loss for a group of approximately 70 boys retested annually for 6 years.

Improvement in IQ by 41 malnourished children brought to a well-nourished state in a children's hospital was reported by Poull (57). The mean change was 10 points of IQ. A control group well nourished at both tests did not change in IQ. Poull attributed the gain in IQ of the experimental group to improvement in nutrition. It is possible, however, that other factors in the children's environment were changed simultaneously.

VII. Recapitulation of Major Issues

1. Comparability of Studies

With respect to the studies of Skeels, Skodak, Wellman, and their associates at Iowa, it may be asked why these discoveries are so late in arriving, or why other research centers have not verified them in detail.

It will be noted from the studies reviewed in this chapter that the Iowa workers can in no sense lay claim, either theoretically or practically, to the sole discovery of substantial and durable environmental effects upon intelligence. Binet, long ago, expounded the theory, and many workers subsequently have reported similar findings.

The Skeels-Skodak data, especially, afford a good touchstone of the comparability of other data and studies over the country. Before anyone asserts that similar results have not been found elsewhere, he should ask of the investigator whose work may have been of a comparable sort these questions:

1. Were the children taken from their true mothers in infancy; that is, in the first few months of life?
2. Were they placed, not in institutions or boarding homes, nor in unselected foster homes, but in socially selected foster homes *for adoption?*
3. Were the children followed for as long as 4 or 5 years?
4. Were individual Binet's used for all tests?
5. Was the sampling substantial, say at least 50 cases?
6. Was reliable case information available on the mothers and fathers, as a check upon the results of adult mental testing?
7. Were any eliminations from the sampling strictly medical or legal in character; and were all eliminations accounted for?

One may turn the main question the other way around: What investigators, whose data are consistent with the implications of the 7 questions, have reported anything incompatible with the Skeels-Skodak findings?

A similar pattern may be built up to serve as a criterion of comparability with the Davenport orphanage experiment thus: What workers, other than the Iowa group, have placed a modern preschool in a standard orphanage and sorted the orphanage population into two well-matched groups, one to receive preschool training up to as much as 400 days, and the other to receive none at all? Similarly, for studies of preschool and elementary-school impacts. In most respects the Iowa data are based upon a more substantial sampling, and one studied over a longer period of time, than other studies now available. As new data are reported elsewhere, it may be that the Iowa results will not be regarded as exceptional.

The materials reported by R. L. Thorndike and others in Part II, Chapter XXIV, under the title "Retest changes in the IQ in certain superior schools," illustrate this point. It will be noted that three schools were investigated: The Ethical Culture School, the Horace Mann School, and the Lincoln School. Binet retests were given at an interval of at least 2½ years. It was found that in two of these schools there was a slight, but insignificant change upward; but that in a third school, with a population of 404 elementary-school children, there was a significant mean gain of 6.17 points of IQ. For the present writers the results in this school have another significance: they are consistent with our findings in the University of Iowa Elementary School, where Wellman has reported a mean gain of 5.6 points. Had Thorndike and his associates chanced to measure only the first two schools, they could have reported that in these fine New York schools there was nothing to support the Iowa findings. But it so happens that one of the schools does lend support to the Iowa findings. Perhaps when good schools over the country tabulate changes in Binet IQ's over a period of years, they will find similar changes, and others of them may be consistently and significantly upward.

There are other interesting data in the Thorndike chapter referred to above. In 1937, Psyche Cattell (10) reported on 3,331 Stanford-Binet IQ changes based on data from the Harvard Growth Study. She showed that about one-tenth of 1 percent varied over 40 points in IQ; 1 percent gained over 30 points in IQ; 5 percent gained over 19 points; 10 percent gained over 14 points; and 25 percent gained over 7 points. On the other hand, 1 percent lost over 16 points; 5 percent over 14 points; 10 percent over 10 points; and 25 percent over 5 points. But in Thorndike and others, on the basis of 1,166 cases in the three schools, 1 percent gained over 40 points; 3.3 percent gained over 30 points;

12.3 percent gained over 20 points; 21.3 percent gained over 15 points; and 49.8 percent over 5 points. Turning to losses in IQ reported by Thorndike, one finds that 1 percent lost over 30 IQ points; 3.5 percent lost over 20 IQ points; 10.4 percent lost over 15 points; 20.9 percent lost over 10 points, and 35.2 percent lost over 5 points.

It will be noted that in the Thorndike data, which are derived from a population already highly selected (the mean IQ's for the three schools were respectively 118.6, 117.5, and 118.8), no known factors of attenuation, regression, or faulty standardization of the Stanford Revision could operate to produce this large excess of gains over losses.

These highly selected New York children — selected also in terms of socio-economic status — showed changes of as much as 20 IQ points in 15.8 percent of the cases. Such data should prepare the mind for changes of even greater magnitude when home and school conditions are allowed to vary from extreme impoverishment to extreme enrichment. And they should give pause to those psychologists who still adhere to the concept of a constant IQ.

2. The Validity of Testing

From time to time studies in this field, especially those involving matched groups or experimental situations, are criticized from the standpoint of a possible lack of rapport in the individual mental testing. So far as Iowa workers are concerned, it may be pointed out that it has long been a tradition in the Iowa Child Welfare Research Station to pay close attention to this factor. It is doubtful whether the first-hand testing experience of some of the Station workers could be surpassed anywhere. Research assistants, already highly selected for ability and previous accomplishment, are given careful training in testing procedures.

It should be pointed out, too, that any departures from good rapport or good testing technique serve to diminish the differences found in experimental situations. If the tests were completely unreliable, or if they were administered in haphazard fashion, we could hardly expect to find any significant differences between experimental and control groups or any constant tendencies toward gain or loss in IQ. Better standardization of the original Binet tests and better testing procedures would serve to emphasize the validity of the results.

There has been a question, too, about the utilization of the Kuhlmann Revision of the Binet tests in the early preschool ages. It is the judgment of the Iowa workers, shared by many others, that the Kuhlmann Revision is the best test below 3½ years. Had the 1916 Stanford Revision been used for this age period, there would have been strong criticism on that account. It may be that the new Stanford Revision will prove equally effective from age two onward. Presumably comparative data will shortly be made available in this connection.

In the meantime it seems worth while to point out that those persons who ascribe great weight to hereditary factors, with the thought that intelligence is primarily innate, should insist upon earlier testing of children, on the ground that the younger the child, the less will be the possible influence of environment, and the more valid the evidences of heredity.

Theoretically, if testing could be done when the child is born, we should have a piling up of hereditary factors, with only those environmental influences present that were mediated through nutritional, mechanical, or chemical factors during the pregnancy of the mother. But, as things are, at birth no behavior items and few anatomical or physical items — short of definitely pathological malformations — give any hint as to the future mental development of the child. Heredity, viewed as a composite of all factors carried in the genes, has transmitted its full influence, but the results are as yet indeterminate. That, later on, some genetically conditioned growth and behavior patterns will arise in the organism, under ordinary environmental circumstances, no one can deny. Nevertheless, as the child grows older, the validity of the genetic factor, as such, in *behavior* situations becomes increasingly difficult to establish.

In any case, those who hold that the impact of environment is negligible must keep close to a zero line of relationship, even when tests are corrected for attenuation and other deficiencies. Their case can be clean-cut only if significant changes cannot be established. If heredity were really as potent as some workers imply, it would make very little difference what kind of environment the child experienced. To illustrate: the point is often made that canal-boat, slum, or backwoods families (children as well as parents) are found in these places because of their inherent mental defectiveness, not because these poor environments are depressing to mental ability.

This points the way to a crucial, large-scale experiment in which the children are taken out of such environments, at birth or shortly thereafter, placed in excellent private homes and there, under conditions of affection and insight, offered standard upper-class social, educational, and vocational opportunities. Hereditary factors should survive this test: if heredity is all-powerful, these children would still be almost as dull as their true parents. But where, the whole country over, during the past 30 years, has this actually been done with a good sampling? For partial answer we can turn to such work as that of Sherman on mountain children; of Klineberg on migrating Negroes; and of Skeels, whose systematic analysis of the mental development of foster children comes closest to meeting the essential criteria.

3. Heredity and Environment

Even when changes are shown to be substantial in amount and duration, the organism itself need never be ruled out of the picture. After all, the child can only be what he could have become. Any mental stature that he attains must be consistent with his heredity, his organic constitution, and his environment. In this sense, since all behavior items reflecting mental growth or mental change are mediated through the organism itself, there need be nothing mysterious about the combination of internal and external factors. In the life space of the child these two are never separated: he is a flexible, changeable, responding organism within wide limits set by heredity and other organic conditions, and within other wide limits set by environmental stimulations and opportunities.

There is much loose talk, chiefly by analogy, concerning the genetic characteristics of human behavior. It is not uncommon for persons to assign family resemblances in such matters as intelligence, learning, music, and literary ability almost exclusively to hereditary factors. That children inherit not only their organism, but also their dynamic social environment, is obvious. It should be obvious, too, that for a given amount of talent in a child, say in music, parents who are musically able or musically interested will seize upon this talent for much encouragement and education. Other children, equally gifted in music, but not endowed with musical families, will be allowed to languish. They may even be severely discouraged.

Equally hazardous are the strange analogies between the highest forms of human behavior, as expressed in learning and intelligence, and the lower forms of growth in species like the fruit-fly. Masses of data on the fruit-fly, gathered under standardized conditions over countless fly generations, do indeed establish a mechanism of heredity for the physical aspects of this invaluable creature. Thus, according to Morgan, 50 genetic factors establish his eye color, 15 his body color, and 10 the length of his wing.

Eventually we may expect scientific evidence and concomitant theory in genetics to reach beyond physical structure and physiological function to the simpler behavior, perhaps reflexive in character, of the lower organisms. How far it may reach into the complex behavior, say intelligence, of a complex animal such as man, no one can foretell. In addition to man's slow rate of generation and maturation and his deeply rooted tradition against controlled propagation, there stands the formidable barrier of a thoroughly mixed genetic constitution. And there is a geometric difficulty arising out of two two-dozen sets of chromosomes and millions of genes.

However, anyone will admit that the talented along any line inherit appropriate potentialities. The very fact that they are talented indicates, in a circular fashion, that they could have been talented. But this does not mean that their parents were talented or that their children will be talented; nor does it indicate what proportion of their talent was ascribable to hereditary factors.

Perhaps mental ability can be hung on some such framework as this:

a. Hereditary factors (those carried in the genes)
b. Constitutional factors (characterizing a specific individual, but not carried in genes and not a resultant from social pressures; *e.g.*, prenatal hemorrhages)
c. Environmental factors (nutrition, stimulation, and education)

One may regard intelligence as always an emergent from these three factors, recognizing that one of these factors may vary from almost negligible to almost crucial strengths. Thus, any gene-carried tendency toward microcephaly may result in feeblemindedness regardless of the influences of *b* and *c*, but potential genius may be unlucky in *b* and pathologically depressed in *c*, so as never to emerge in the individual. As one goes toward the central tendencies of the distributions of these factors, results are obtained that are complex, unpredictable, and perhaps truly indeterminate.

Certainly one cannot solve the practical social problem of feeblemindedness or of genius by paying attention only to these extremes found in the general distribution of the population. Jennings (30) estimates that, if one were to assume that a single gene could carry a feeblemindedness taint, there would be 30 times as many individuals with this gene as there would be feebleminded persons. Thus, to a geneticist, the mass of feeblemindedness, as well as of genius, arises out of the mass of the people.

The intricate nature of the problem, whether in the fruit-fly or the human being, is indicated by Jennings (30):

> Under certain conditions, therefore, a given set of genes may produce a defective individual; under other conditions, a normal one. The defectiveness of the individuals could therefore be prevented either by starting them with a different set of genes, or by changing the environment in which they develop. In the fruit-fly, . . . defectiveness of a certain gene produces under ordinary conditions an abnormal structure of the abdomen. But if the individuals are kept in a dry atmosphere, the abnormality does not appear. A change of environment has prevented the effects of a defective gene. . . .
>
> Similarly, in man, certain combinations of genes yield an individual that is unusually susceptible to infection with tuberculosis. But if infection is prevented, or if the resistance of the individual is increased by proper living conditions or by therapeutic measures, he does not succumb to tuberculosis. Such prevention, by environmental means, of the consequences of weak or imperfect genes plays a very great rôle in all matters of susceptibility and resistance. Essentially the same in general principle is the effect of education, training, and the like, in changing the behavior that would without it result from the genes present in the individual. In some of these matters the effects of genes and environment are inextricably intermingled. Behavior is bound to be relative to environment, it cannot be dealt with as dependent on genes alone. A given set of genes may result under one environment in criminality; under another in the career of a useful citizen. (p. 229) [1]

It may be asked, with respect to any studies showing substantial environmental effects upon mental growth: Has the hereditary endowment of these children been modified? Perhaps the answer is that no one knows. What contribution to his intelligence does a child get from heredity, anyway? It is well known that heredity does not presuppose likeness; it may also produce unlikeness. But, the objector argues, surely you cannot turn feebleminded children into normal ones.

Perhaps this is a good place to point out certain theoretical difficulties in which an authority like Tredgold (80) has placed himself. To quote:

> In a certain proportion of cases the mental defect of the individual has been preceded by a similar defect in his ancestors, and it may then be regarded as a definite instance of 'heredity' in the strict meaning of the word. But this is by no means always the case; indeed, it is not even

[1] Quoted from *The Biological Basis of Human Nature,* by permission of W. W. Norton, publisher, New York, 1930. See Reference 30.

> the rule, and in my experience it is commoner for ancestors of defectives to suffer from such conditions as insanity, epilepsy, dementia, and allied psychopathological states than it is for them to be actually mentally deficient. It is obvious that the repeated occurrence of such conditions as these in a family, the members of which live in an environment which does not differ essentially from that of the mentally normal, must be indicative of some abnormality of the germ plasm; in other words, of a germinal variation. (pp. 24–25) [1]

In short, according to Tredgold, the inheritance of mental defect, which is so freely assumed by some psychologists, is more often an inheritance running from insanity or epilepsy in the parent to mental defect in the child. Is it not true that many psychologists, relying upon a hereditary chain of events, have been thinking exclusively of a line of mental defect in the parent running to mental defect in the child?

Moreover, one might question Tredgold's use of the word *essentially* in the foregoing quotation: perhaps a family marked by insanity or epilepsy does differ, in its environmental impacts upon the child, from families not similarly handicapped.

Later on, Tredgold encounters a real dilemma:

> Now the essence of mental defect is that it is incurable; when, therefore, it is found that defectives attending special schools are returned 'cured' to the ordinary schools, it is clear that very few, if any, of them were really defectives according to the Mental Deficiency Act. My inquiries showed that in various towns in England the proportion so returned as cured varied from 2 to as much as 20 percent, while Dr. Eich-

---

[1] In the Sixth Edition (1937) of *Mental Deficiency,* Tredgold states his position in somewhat different terms, but the meaning is similar. For example:

> There is, in my opinion, no doubt that there is an exceptional concentration of various forms of mental abnormality in the families of the great majority of defectives (p. 28). . . . In the families of some defectives there is no antecedent defect at all, and in most of them there is a combination of various mental abnormalities such as insanity, psychoneurosis, epilepsy, dementia, and dullness (p. 31). . . . Again, whilst it is perfectly true that many defectives come of intellectually dull stocks, there are many who come of families of average and even superior ability, but numerous members of which have been insane or otherwise psychopathic (p. 32). . . . We can, I think, say with certainty that the germ cells are far from being immutable, and that alterations in them can be produced by various factors of the environment (p. 38).

Quotations made by permission of the Williams & Wilkins Company, American agents for the British publishers.

> holz, H.M. Inspector of Special Schools, in giving his evidence before the Royal Commission, stated that according to his observations from 40 to 50 percent of the cases classed as feebleminded recovered under care and training. (p. 180) [1]

According to Tredgold, then, one can never say that a child is mentally defective, since at some time he may grow out of it. If he grows out of it, then he was never mentally defective in the first place! Hence all valid diagnoses of mental deficiency would have to be post-mortems.

To the present writers, Tredgold's earlier position was far-fetched and sterile. If, on the basis of reliable individual testing, supported by clinical evidence in terms of defective behavior, learning, and achievement, a child is judged feebleminded, the judgment appears sound, even if subsequently he changes his status. That any child could be labeled retroactively as never having been defective would seem to beg the question.

4. Utilization of Certain Concepts

It is sometimes held that increases in IQ related to improved environment, for example, in certain nursery schools or elementary schools, cannot be valid unless the improvement is not only significant in the early period of attendance, but also is maintained progressively. But the first upward turn may result from a shifting to a relatively enriched environment. To postulate steady acceleration thereafter, one has to assume that the school's program is adjusted to accelerating children. In our experience such a program is rarely available. It has not been practicable, for example, in the Iowa laboratories to set up programs that are adequately stimulating to those initially very bright, or who become very bright. It may be that schools should turn to a tutorial practice in which each very bright child receives attention through the assignment of tasks and experiences consistent with his ability.

The phenomenon of *regression* to explain (or explain away) results should be employed with caution. On retesting with tests not completely reliable, children of high or low IQ will tend to move toward the

---

[1] In the Sixth Edition (1937), Tredgold does not cite Eichholz, but does adhere to his earlier position. For example: "But *cure* [of the mentally deficient] will never take place (p. 453). . . . No means exist, or ever will exist, by which we can *supply* intelligence to the mentally deficient" (p. 461).

Citations are made from the 1922 Edition because they have entered into the thinking and writing of many psychologists over the past fifteen years.

mean of the second testing; those initially high will be less high; those initially low, less low. Regression cannot explain a mass upward or downward shift of a whole population, as in the Wellman preschool gains in IQ, or the superiority of Skeels' foster-home children. Nor can all upward or downward shifting of individuals within a given population be ascribed automatically to this phenomenon.

Another concept that should not be used carelessly is *selection.* Selection, properly utilized as a statistical or experimental concept, refers to the error that results from an unknown or uncontrolled variable affecting in turn one or more variables being analyzed in such a way as to vitiate the results. For example, if one were comparing the mental abilities of children in various occupational classes and had, in a laboring class, a selection of the brightest children and, in a professional class, a selection of the dullest children, he might conclude that children of laboring-class families are brighter than children of the professional groups. Selection here would color his results and damage his conclusions.

But one may have a selection of cases not related to any of the variables under consideration. For instance, the selection of a small number of bright children in the preschool or the early elementary school out of a larger population of similarly bright children need not produce distortion as between the small sampling and the large. There is still a sampling problem, and the burden of proof is on the investigator to show that his sample is representative.

In short, it is evident that *criticism* of studies, as well as their planning and execution, has necessary tenets of validity. Some difficulties, as Dr. Freeman has pointed out in this Yearbook, arise out of differences in definition, or perhaps in methodology.

In essence what may be termed the 'Iowa-Binet theory of intelligence' simply permits a large amount of change in a child's brightness through environmental impingements on the organism: the growing child changes his rate of growth. In any given study this change may be slight. But where the change in rate (that is, the shift in intelligence quotient) is substantial, and demonstrable for a large sampling, no amount of reference to other studies can gainsay it.

Nevertheless, the scientific requirement of eventual verification of results and conclusions by other research centers is a good one. The Iowa workers share the common hope that all theories and conclusions will be rigorously subjected to this test.

References

(1) Arthur, G. "The predictive value of the Kuhlmann-Binet Scale for inmates of a state school for the feebleminded." *Jour. Applied Psychol.,* 17: 1933, 188–194.

(2) Asher, E. J. "The inadequacy of current intelligence tests for testing Kentucky mountain children." *Ped. Sem. and Jour. Genet. Psychol.,* 46: 1935, 480–486.

(3) Barrett, H. E., and Koch, H. L. "The effect of nursery-school training upon the mental-test performance of a group of orphanage children." *Ped. Sem. and Jour. Genet. Psychol.,* 37: 1930, 102–122.

(4) Bayley, N., and Jones, H. E. "Environmental correlates of mental and motor development: A cumulative study from infancy to six years." *Child Develop.,* 8: 1937, 329–341.

(5) Beckham, A. S. "A study of the intelligence of colored adolescents of different social-economic status in typical metropolitan areas." *Jour. Soc. Psychol.,* 4: 1933, 70–91.

(6) Bird, G. E. "Effect of nursery-school attendance upon the mental growth of children." (Part II, Chapter IV, of this *Yearbook*)

(7) Bradway, K. P. "Paternal occupational intelligence and mental deficiency." *Jour. Applied Psychol.,* 19: 1935, 527–542.

(8) Burks, B. S. "The relative influence of nature and nurture upon mental development: A comparative study of foster parent-foster child resemblance and true parent-true child resemblance." [In] *The Twenty-Seventh Yearbook, Part I,* of this Society, 1928, pp. 219–316.

(9) Byrns, R., and Henmon, V. A. C. "Parental occupation and mental ability." *Jour. Educ. Psychol.,* 27: 1936, 284–291.

(10) Cattell, P. "Stanford-Binet IQ variations." *School and Soc.,* 45: 1937, 615–618.

(11) Chipman, C. E. "The constancy of the intelligence quotient of mental defectives." *Psychol. Clin.,* 18: 1929–1930, 103–111.

(12) Coffey, H. S. *A Study of Certain Mental Functions and Their Relation to Changes in the Intelligence of Preschool Children.* (State University of Iowa, Unpublished Doctoral dissertation, 1938, 161 ms. pp.)

(13) Coffey, H. S., and Wellman, B. L. "The rôle of cultural status in intelligence changes of preschool children." *Jour. Exper. Educ.,* 5: 1936–1937, 191–202.

(14) Crissey, O. L. *Mental Development as Related to Institutional Residence and Educational Achievement.* (University of Iowa Studies in Child Welfare, Vol. 13, No. 1: Iowa City, 1937) 81 pp.

(15) Davenport, C. B., and Minogue, B. M. "The intelligence quotient and the physical quotient: Their fluctuation and intercorrelation." *Human Biol.,* 2: 1930, 473–507.

(16) Dubnoff, B. "A comparative study of mental development in infancy." *Ped. Sem. and Jour. Genet. Psychol.,* 53: 1938, 67–73.

(17) Edwards, A. S., and Jones, L. "An experimental and field study of north Georgia mountaineers." *Jour. Soc. Psychol.,* 9: 1938, 317–333.

(18) FRANDSEN, A., and BARLOW, F. P. "Influence of the nursery school on mental growth." (Part II, Chapter VII, of this *Yearbook*)

(19) FREEMAN, F. N., HOLZINGER, K. J., MITCHELL, B. C., and others. "The influence of environment on the intelligence, school achievement, and conduct of foster children." [In] *The Twenty-Seventh Yearbook, Part I,* of this Society, 1928, pp. 103–217.

(20) FURFEY, P. H. "The relation between socio-economic status and intelligence of young infants as measured by the Linfert-Hierholzer scale." *Ped. Sem. and Jour. Genet. Psychol.,* 35: 1928, 478–480.

(21) FURFEY, P. H., and MUEHLENBEIN, J. "The validity of infant intelligence tests." *Ped. Sem. and Jour. Genet. Psychol.,* 40: 1932, 219–223.

(22) GOODENOUGH, F. L. "The relation of the intelligence of preschool children to the education of their parents." *School and Soc.,* 26: 1927, 54–56.

(23) GOODENOUGH, F. L. "The relation of the intelligence of preschool children to the occupation of their fathers." *Amer. Jour. Psychol.,* 40: 1928, 284–302.

(24) GOODENOUGH, F. L., and MAURER, K. M. "The mental development of nursery-school children compared with that of children without nursery-school training." (Part II, Chapter IX, of this *Yearbook*)

(25) HILDEN, A. H. "A comparative study of the intelligence quotient and Heinis' personal constant." *Jour. Applied Psychol.,* 17: 1933, 355–375.

(26) HILDRETH, G. "Occupational status and intelligence." *Personnel Jour.,* 13: 1934, 153–157.

(27) HILDRETH, G. "Adopted children in a private school." (Part II, Chapter X, of this *Yearbook*)

(28) HIRSCH, N. D. M. "An experimental study of east Kentucky mountaineers: A study in heredity and environment." *Genet. Psychol. Monog.,* 3: 1928, 189–244.

(29) HOLZINGER, K. J. "Reply to special review of 'Twins.'" *Psychol. Bull.,* 35: 1938, 436–444.

(30) JENNINGS, H. S. *The Biological Basis of Human Nature.* (W. W. Norton: New York, 1930) 384 pp.

(31) JONES, H. E., CONRAD, H. S., and BLANCHARD, M. B. "Environmental handicap in mental-test performance." *Univ. Calif. Pub. Psychol.,* 5, No. 3: 1932, 63–99.

(32) JONES, H. E., and JORGENSON, A. P. "Mental growth as related to nursery-school attendance." (Part II, Chapter XII, of this *Yearbook*)

(33) JORDAN, A. M. "Parental occupations and children's intelligence scores." *Jour. Applied Psychol.,* 17: 1933, 103–119.

(34) KAWIN, E. *Children of Preschool Age: Studies in Socio-Economic Status, Social Adjustment, and Mental Ability, with Illustrative Cases.* (University of Chicago Press: Chicago, 1934) 340 pp.

(35) KAWIN, E., and HOEFER, C. *A Comparative Study of a Nursery-School versus a Non-Nursery-School Group.* (University of Chicago Press: Chicago, 1931) 52 pp.

(36) KEPHART, N. C. "Possibilities for influencing the rate of mental growth in retarded children through environmental stimulation." (Part II, Chapter XIII, of this *Yearbook*)

(37) KLINEBERG, O. *Negro Intelligence and Selective Migration.* (Columbia University Press: New York, 1935) 66 pp.

(38) KOUNIN, J. *The Effect of Preschool Attendance upon Later School Achievement.* (State University of Iowa, Unpublished Master's Thesis: Iowa City, 1939) 92 ms. pp.

(39) KUHLMANN, F. "The results of repeated mental reëxaminations of 639 feebleminded over a period of ten years." *Jour. Applied Psychol.,* 5: 1921, 195–224.

(40) LAMSON, E. E. "To what extent are intelligence quotients increased by children who participate in a rich, vital school curriculum?" *Jour. Educ. Psychol.,* 29: 1938, 67–70.

(41) LAMSON, E. E. "A follow-up study of a group of nursery-school children." (Part II, Chapter XIV, of this *Yearbook*)

(42) LAWRENCE, E. M. "An investigation into the relation between intelligence and inheritance." *Brit. Jour. Psychol. Monog. Suppl.,* 16, No. 5: 1931, 80 pp.

(43) LEAHY, A. M. "Nature-nurture and intelligence." *Genet. Psychol. Monog.,* 17: 1935, 235–303.

(44) LICHTENSTEIN, M., and BROWN, A. W. "Intelligence and achievement of children in a delinquency area." *Jour. Juv. Research,* 22: 1938, 1–25.

(45) LITHAUER, D. B., and KLINEBERG, O. "A study of the variation in IQ of a group of dependent children in institution and foster home." *Ped. Sem. and Jour. Genet. Psychol.,* 42: 1933, 236–242.

(46) LORIMER, P., and OSBORN, F. *Dynamics of Population; Social and Biological Significance of Changing Birth Rates in the United States.* (Macmillan: New York, 1934) 461 pp.

(47) MCALPIN, A. S. "Changes in the intelligence quotients of Negro children." *Jour. Negro Educ.,* 1: 1932, 44–48.

(48) MCNEMAR, Q. "Special review of Newman, Freeman, and Holzinger's twins: A study of heredity and environment." *Psychol. Bull.,* 35: 1938, 237–249.

(49) MCNEMAR, Q. "Rejoinder to Holzinger's reply to special review of 'Twins,'" *Psychol. Bull.,* 35: 1938, 552–554.

(50) MCCANDLESS, B. R. *The Effect of Enriched Educational Experiences upon the Growth of Intelligence of Very Superior Preschool Children.* (State University of Iowa, Unpublished Master's Thesis: Iowa City, 1938) 127 ms. pp.

(51) National Society for the Study of Education. *The Twenty-Seventh Yearbook. Nature and Nurture, Part I: Their Influence Upon Intelligence.* (Public School Publishing Co.: Bloomington, Ill., 1928) 465 pp.

(52) NEFF, W. S. "Socio-economic status and intelligence: A critical survey." *Psychol. Bull.,* 35: 1938, 727–757.

(53) NEWMAN, H. H., FREEMAN, F. N., and HOLZINGER, K. J. *Twins: A Study of Heredity and Environment.* (University of Chicago Press: Chicago, 1937) 369 pp.

(54) OLSON, W. C., and HUGHES, B. O. "Subsequent growth of children with and without nursery-school experience." (Part II, Chapter XV, of this *Yearbook*)

(55) OTIS, M. "Improvement of feebleminded girls over 16 years of age." *Jour. Applied Psychol.*, 13: 1929, 301–315.

(56) PATERSON, D. G., and RUNDQUIST, E. A. "The occupational background of feeblemindedness." *Amer. Jour. Psychol.*, 45: 1933, 118–124.

(57) POULL, L. E. "The effect of improvement in nutrition on the mental capacity of young children." *Child Develop.*, 9: 1938, 123–126.

(58) PRITCHARD, M. C., HORAN, K. M., and HOLLINGWORTH, L. S. "The course of mental development in slow learners under an 'experience curriculum.'" (Part II, Chapter XVI, of this *Yearbook*)

(59) PROUTY, R. A. "Psychological classification versus clinical diagnosis." *Psychol. Clin.*, 18: 1929–1930, 213–220.

(60) REYMERT, M. L., and HINTON, R. T. "The effect of a change to a relatively superior environment upon the IQ's of one hundred children." (Part II, Chapter XVII, of this *Yearbook*)

(61) RISDEN, G. A. *An Exploratory Study in Education for Slowly Developing Children.* (Ohio State University Abstracts of Doctor's Dissertations, No. 24: Columbus, Ohio, 1937) pp. 279–293.

(62) ROBERTSON, J. *A Study of the IQ Changes of Children of 120 IQ and above on Entering School.* (State University of Iowa, Unpublished Master's Thesis: Iowa City, 1936) 105 ms. pp.

(63) RUSK, R. R. "The basic intelligence of Scottish children." (Part II, Chapter XVIII, of this *Yearbook*)

(64) SCHOTT, E. L. "IQ changes in foster-home children." *Jour. Applied Psychol.*, 21: 1937, 107–112.

(65) SHERMAN, M., and KEY, C. B. "The intelligence of isolated mountain children." *Child Develop.*, 3: 1932, 279–290.

(66) SKEELS, H. M. "Mental development of children in foster homes." *Jour. Consulting Psychol.*, 2: 1938, 33–43.

(67) SKEELS, H. M. "Some Iowa studies on the mental growth of children in relation to differentials in environment: A summary." (Part II, Chapter XX, of this *Yearbook*)

(68) SKEELS, H. M., and FILLMORE, E. A. "Mental development of children from underprivileged homes." *Ped. Sem. and Jour. Genet. Psychol.*, 50: 1937, 427–439.

(69) SKEELS, H. M., UPDEGRAFF, R., WELLMAN, B. L., and WILLIAMS, H. M. *A Study of Environmental Stimulation: An Orphanage Preschool Project.* (University of Iowa Studies in Child Welfare: Vol. 15, No. 4: Iowa City, 1938) 191 pp.

(70) SKODAK, M. "The mental development of adopted children whose true mothers are feebleminded." *Child Develop.*, 9: 1938, 303–308.

(71) SKODAK, M. *Children in Foster Homes: A Study of Mental Development.* (University of Iowa Studies in Child Welfare, Vol. 16, No. 1: Iowa City, 1939) 156 pp.

(72) SNYGG, D. "The relation between the intelligence of mothers and of their children living in foster homes." *Ped. Sem. and Jour. Genet. Psychol.*, 52: 1938, 401–406.

(73) SPEER, G. S. "The mental development of children of feebleminded and normal mothers." (Part II, Chapter XXI, of this *Yearbook*)

(74) STARKWEATHER, E. K. *IQ Changes over a Long Interval in Relation to Sex and Group Mental Level.* (State University of Iowa, Unpublished Master's Thesis: Iowa City, 1938) 79 ms. pp.

(75) STARKWEATHER, E. K., and ROBERTS, K. E. "IQ changes occurring during nursery-school attendance at the Merrill-Palmer School." (Part II, Chapter XX, of this *Yearbook*)

(76) TERMAN, L. M., and MERRILL, M. A. *Measuring Intelligence: A Guide to the Administration of the New Revised Stanford-Binet Tests of Intelligence.* (Houghton Mifflin: Boston, Mass., 1937) 461 pp.

(77) THORNDIKE, R. L. "Note on 'IQ changes in foster-home children' by Emmett L. Schott." *Jour. Applied Psychol.*, 22: 1938, 662.

(78) THORNDIKE, R. L. "Retest changes in the IQ in certain superior schools." (Part II, Chapter XXIV, of this *Yearbook*)

(79) TRAINOR, J. C. "Experimental results of training in general semantics upon intelligence-test scores." [In] *Papers from the First American Congress for General Semantics,* held at Ellensburg, Washington, March 1 and 2, 1935, 58–67.

(80) TREDGOLD, A. F. *Mental Deficiency.* (William Wood & Co.: New York, 1927) 569 pp.

(81) UPDEGRAFF, R., DAWE, H. C., FALES, E. E., STORMES, B. E., and OLIVER, M. G. *Practice in Preschool Education.* (McGraw-Hill: New York, 1938) 408 pp.

(82) VAN ALSTYNE, D. *The Environment of Three-Year-Old Children: Factors Related to Intelligence and Vocabulary Tests.* (Teachers College, Columbia University, Contributions to Education, No. 366: New York, 1921) 108 pp.

(83) VOAS, W. A. "Does attendance at the Winnetka Nursery School tend to raise the IQ?" (Part II, Chapter XXV, of this *Yearbook*)

(84) WARING, E. B. "A report of the psychological examinations of preschool children in the nursery school over a period of eight years." [In] "Ten-Year Report of Studies in Child Development and Parent Education," *Cornell Univ. Agric. Experiment Station, Cont. from Stud. in Home Econ. Bull.*, No. 638: 1935, 42–43.

(85) WELLMAN, B. L. "Some new bases for interpretation of the IQ." *Ped. Sem. and Jour. Genet Psychol.*, 41: 1932, 116–126.

(86) WELLMAN, B. L. "The effect of preschool attendance upon the IQ." *Jour. Exper. Educ.*, 1: 1932–1933, 48–69.

(87) WELLMAN, B. L. "Growth in intelligence under differing school environments." *Jour. Exper. Educ.*, 3: 1934–1935, 69–83.

(88) WELLMAN, B. L. "Mental growth from preschool to college." *Jour. Exper. Educ.*, 6: 1937–1938, 127–138.

(89) WELLMAN, B. L. *The Intelligence of Preschool Children as Measured by the Merrill-Palmer Scale of Performance Tests.* (University of Iowa Studies in Child Welfare, Vol. 15, No. 3: Iowa City, 1938) 150 pp.

(90) WELLMAN, B. L. "Iowa studies on the effects of schooling." (Part II, Chapter XXVI, of this *Yearbook*)

(91) WELLS, J., and ARTHUR, G. "Effect of foster-home placement on the intelligence ratings of children of feebleminded parents." *Mental Hygiene,* 23: 1939, 277–285.

(92) Wheeler, L. R. "A study of the mental growth of dull children." *Jour. Educ. Psychol.*, 21: 1930, 367–378.

(93) Wheeler, L. R. "The intelligence of east Tennessee mountain children." *Jour. Educ. Psychol.*, 23: 1932, 351–370.

(94) Woodall, C. S. "Analysis of I.Q. variability." *Proc. and Addr. Amer. Assoc. Stud. Feebleminded*, 36: 1931, 247–262.

SECTION V

COMMENTS UPON THE YEARBOOK AS A WHOLE BY MEMBERS OF THE YEARBOOK COMMITTEE

CHAPTER XV

PERSONAL REACTIONS OF THE YEARBOOK COMMITTEE[1]

I

LEONARD CARMICHAEL

Psychology today is properly classified in certain of its aspects as a biological science and in certain of its aspects as a social science. It follows because of this double reference that the problems involved in describing the nature and nurture of intelligence become difficult to solve. Human reactions, including those acts called intelligent, are an expression of the way in which anatomical structures function under specific conditions. Nevertheless, no living human individual develops in a social vacuum, and one of the outstanding characteristics of the structure of the human organism is that its very protoplasmic make-up is modified as a result of processes such as inner growth and also as a result of processes initiated by environmental stimuli. Especially important among the environmental stimuli that determine changes in individuals are other human beings who make up society, social traditions, and social institutions.

As I see it, the major contribution of the present Yearbook is the elaborate discussion that it gives of techniques, some of which had not been developed a decade ago, by means of which intrinsic biological differences between human beings can be to some extent differentiated from purely cultural or socially determined differences. The reader of the Yearbook cannot fail to be impressed by the fact that, still in the year 1939, honest differences of opinion exist on the part of professional students concerning the effect of the environment upon the fundamental modification of the organism in regard to those processes that make intelligent behavior possible. In extreme cases, all serious stu-

[1] On account of temporal limitations, it has been impossible to provide each member of the Committee with corrected proof of the entire Yearbook before it became necessary to request the submission of the individual appraisals that follow. On this account, there may appear statements — quotations, for example — that do not accord precisely with the text of the Yearbook as finally emended by the contributors and by the editor. — *Editor.*

dents agree that the limits of an organism's capacity to perform intelligent acts are determined by structures that exist in the individual as a result of hereditary factors. No one believes that a chimpanzee, even if given all the advantages of a human education, would ever be able to pass an examination in high-school algebra. On the other hand, it is equally clear that 'intelligence' is never fully measured in the normal adult human individual save through the medium of socially acquired knowledge and skills. Just as the chimpanzee cannot learn algebra, no matter what his training may be, so also few believe that a superior human individual could in one generation, if removed from contact with the social inheritance of the past, develop an ability to perform the operations required in formal mathematics, because such actions require the cumulative work of many centuries.

In less extreme cases the conclusion is not so clear. An answer to the question as to whether or not the social environment ever modifies the *fundamental* structure of the individual upon which intelligent acts depend must be secured by the indirect study of socially acquired acts. In the human subject such acts are always learned responses, and so one must base his judgment of the excellence or inferiority of brain protoplasm by watching the performance of acts of skill that can be carried out under certain conditions, such as greater practice, just as well by the less gifted individual as by the more gifted individual. Unfortunately also certain presuppositions concerning the nature of intelligence, certain judgments concerning what constitutes an effective intelligence test, and certain judgments concerning appropriate statistical and experimental procedures are all necessary if one is to describe correctly the *nature* of intelligence, let alone its *nurture*. Hence there is still room for much difference of opinion in the significant field covered by this book. I believe, however, that there is enough evidence presented here to allow the careful reader to arrive at sound conclusions concerning the present status of the subject.

As I have already suggested in the summary of the part of it for which I am personally responsible, the field of this book must always be of great significance for the teacher. Human intelligence, considered realistically and pragmatically, involves among other skills the capacity of individuals to manipulate concepts and use what may be called 'intellectual tools.' Such intellectual instruments come from the past of the race and can be secured alone by a process of education. Thus the formal educator must allow each individual to acquire to the best of his or her inherited ability those skills and concepts that will permit him to use his biologically given brain in a manner that will be most advantageous to himself and to society. In other words, the practical school man should gain from the careful study of this book, not a feeling of discouragement and confusion about intelligence, but rather, a renewed recognition of the importance of considering each individual in each schoolroom as a biologically unique human personality.

It may be because of my biological bias that I do not see how education *as such* can change the fundamental protoplasm of the individual and render him 'more intelligent.' But it does seem clear that without education, no matter what protoplasm a given individual may have, he cannot do well on most intelligence tests or in most life situations in civilization that require intelligent behavior. Thus in a sense the educator may well teach always as if he were aiming to increase intelligence as well as to impart knowledge. Everything in this book suggests that he should not be discouraged, however, if the same educational procedures do not produce the same results in all individuals. By examinations and other techniques the educator selects those whose abilities fit them for advanced education and for the professions requiring quite literally the 'best brains.' Knowledge of the material of this book will assist the educator, whatever his special province may be, to strike a proper balance between his functions as modifier of human beings and as selector of human beings, and it is hard to overemphasize the importance, in the present state of American education, of striking this balance in a just manner.

II

Frank N. Freeman

Anything like a comprehensive evaluation of the reports of research or the discussions of the Yearbook is impossible because of lack of space and lack of opportunity to examine all the contributions in detail and in their final form. It would be difficult even if these limitations were removed. I shall therefore content myself with a brief statement of one consideration basic to any evaluation that may be made, the neglect of which is, I believe, responsible for much of the confusion that characterizes our thinking on the problem of the factors in intelligence.

It must be puzzling to the onlooker, as it has been puzzling to me, to witness the wide divergence in views and the apparent contradiction in evidence on this question. It seems incredible that evidence can seem to be conflicting or that honest and competent scientists can come to such diverse conclusions. It may be true, as some think, that all the competence and honesty is on one side and all the error and fallacious thinking is on the other, but this seems hardly likely. We may make the assumption that there must be some truth and some error on each side and, hence, conclude that the best procedure is to seek a middle-of-the-road compromise in the hope of getting nearer to the truth by this method than if we adhere to either extreme. This is perhaps the best practical course if we cannot find a better, but it is not one with which we can be well content.

I believe the confusion of thought on this question arises from two facts. The first is that the majority of interpreters of the scientific evidence are looking for proof rather than for the most probable explanation. Proof, of course, is the ideal of scientific work. But it is not always attainable. If we demand proof, we are setting up a very rigid standard for the admission of evidence. We would be compelled to exclude much evidence that we would admit if we were merely trying to assess the probability of one or another conclusion.

The second fact is that one set of thinkers starts with the tentative assumption that one of the alternative views is correct and demands proof of the truth of the opposite view; while the other set of thinkers assumes the truth of the opposite hypothesis and demands proof of the one the first group assumes. In other words, each one lays the burden of proof on the other. Hence the criterion of evidence that each group sets up for the opposing view is of a different order than that which it accepts for its own. An example of this laying of the burden of proof is the requirement that the Iowa group show that gains associated with nursery-school attendance or that gains made by foster children are not due to differences in the tests used before and after, to errors in the early tests, or to biased selection of cases. On the other side, the advocates of environmental influence require that the advocates of heredity prove that family resemblance is not due to similarity of environment or that the constancy of the IQ is not due to the uniformity of the environment.

These are perfectly legitimate demands. All possible ambiguities in the interpretation of the data of investigation should be cleared up and proof should be made as complete as possible. But the demands should be as rigid on one side as on the other. There is no plan, it seems to me, for placing the burden of proof on either side. Burden of proof is a forensic, not a scientific, concept. It presumes a controversy and the rightness or truth of one side until the contrary is proved, as in our jurisprudence the prisoner is presumed innocent until he is proved guilty. But in science there is no presumption on one side or the other. Conclusions should be drawn according to the probabilities as indicated by the weight of evidence.

The principle of burden of proof is more often applied unwittingly than explicitly. The scientist is predisposed toward one side and consequently demands stronger evidence in support of the opposite point of view. He is not aware that he is doing so. It is one of the prime requisites of good judgment on scientific questions to weigh the evidence equally and make the same allowances and the same demands on both sides. The sharp division of opinion on the nature-nurture controversy is due, I believe, to the fact that this principle is disregarded. It is good to demand the most rigid proof, but it is a mistake to interpret the results of the investigation on any question in terms of the requirement of the burden of proof.

III

Florence Goodenough

Inasmuch as my general point of view in regard to the Yearbook topic has already been made fairly explicit in Part I, Chapters XI and XII, and since many of the original contributions in Part II have already been commented on therein, it is unnecessary for me to do more than add a few brief notes by way of summary.

Throughout history, civilization has advanced through the discovery of nature's laws and the adaptation of these laws to the development and utilization of our natural resources. But thus far, little progress has been made toward any basic improvement in that greatest of all resources, the mind of man. The modern American child, it is true, has certain advantages when compared with the Athenian child of the fourth century before Christ. The knowledge acquired by mankind in the course of more than two thousand added years of progress has been organized for his convenience and many special devices to facilitate his learning have been worked out. But that his capacity to learn is greater than that of his Athenian forerunner — had the two been possessed of the same opportunities for learning — or that the range of individual differences in learning ability has been materially changed within the period of written history is highly questionable. Of all our natural phenomena, human mentality seems to be among the most obdurate to modification. This is not to say that no means of effecting such modification will ever be found. Indeed, it is just there that the greatest hope for human advancement lies, for were a real improvement in the quality of man's mind to be accomplished, a solution of many, if not most, of our present social ills might fairly be expected. But the long and unsuccessful search for a means of bringing about such a change makes it seem unlikely that the method, when and if it be found, will be a simple or external one.

Some progress in this direction has already been made. A century ago the cretin was condemned to a life of hopeless idiocy. Now that we know the cause of his defect, we are able at least to ameliorate his condition, in some cases to restore him to normality. When greater knowledge of the physiological factors that affect the mentality of man in general has been gained, we may perhaps hope for the discovery of procedures by which the basic quality of those behavioral capacities that we subsume under the name of 'intelligence' may truly be changed. In the meantime, let us, by all means, strive to provide for every child the maximal opportunity to develop those gifts with which nature has endowed him.

Perhaps, after all, the question of the improvability of intelligence is largely a matter of definition. That it is possible to improve the quality of individual adjustment to modern social demands there can be no reasonable doubt. Other things being equal, the man who can read has a better chance of success in the modern world than the one who is illiterate; one who can make change quickly and correctly has an advantage over one who is unfamiliar with our monetary system; one who is acquainted with ordinary principles of mechanics has a better chance of success than another of equal capacity to learn but of little experience with machinery. In like manner, the person who is reasonably self-confident and has at least a fair degree of social poise is better equipped to get along in the modern world than is the one who suffers from excessive timidity and self-consciousness. Education and training can aid the individual in the acquisition of skills and knowledge and help him to develop suitable attitudes toward himself and the world about him. In this way he may be caused to act more effectively and thus, in a sense, more intelligently. To those who identify 'intelligence' with 'achievement' there can be no question but that the environmental influence upon intelligence, as thus defined, is very great. For those who prefer to think of 'intelligence' as that quality of mind which sets limits to achievement — that is to say, which makes it possible for an individual to progress thus far and no farther under a given set of external conditions — the question is very different. It is well known that under the careful and sympathetic training given in the best of our institutions for the feebleminded, the moron is taught many skills that under the more difficult conditions of the outside world he would be unlikely to acquire. This, however, does not mean that his basic capacity to learn has been improved, but only that the method of teaching has been adapted to his pattern and tempo of learning. Neither does it mean that all barriers to the attainment of skills acquired by his better-endowed fellows can be removed by training. Apart from all other questions, there is the inescapable factor of time to be considered. Even under the most careful tutelage, the moron's span of life is hardly great enough for him to arrive at the point where he would be ready to begin the calculus. Suitable education can aid the individual in realizing his potentialities, but that these potentialities can be materially advanced by training alone has thus far not been demonstrated.

IV

Leta S. Hollingworth

The materials presented in this Yearbook are on the whole of good quality. Perhaps their chief merit would be best described by the adjective 'provocative.' The discussions in Part I provoke intensive critical thinking on the issues involved in the nature-nurture controversy; the original materials in Part II are conducive to further study, since no final agreement is found among the results presented.

The really most important advance represented by this Yearbook appears to be that herein are recorded for the first time the results of *long-time, follow-up studies* on the same individuals, year after year. The studies of gifted children reported by Terman, by Witty, and by Hollingworth enable us for the first time to know what may be expected in the course of development from children who make high scores on intelligence tests in childhood. The studies by Jones, by Wellman, and by many others suggest what may be expected in the course of development from children more or less unselected.

These long-time studies of the predictability of developmental status lean heavily toward the same conclusion: *Existing methods of measuring intelligence have small predictive value when applied before seven or eight years of age; when applied at and after seven or eight years of age, the methods even now available have high predictive power.*

The reasons why tests made at preschool ages do not predict well are various, and have been fully pointed out in the literature reviewed in the Yearbook. Goodenough's critique, Chapter XII of Part I, covers these reasons well. It seems to me that the chief reasons why IQ's change so much when taken at two, three, and four years of age may be stated briefly under three headings: (1) available infant tests are based on samples selected toward the high end of the scale, so that when tested at school age, by tests standardized on more nearly random samples (obtained under compulsory education) the later IQ's tend *to rise* because of the difference in standards; (2) negativism as a phenomenon of early childhood tends strongly in many cases to depress the obtained IQ below the true IQ; (3) children under school age have not yet lived long enough, nor have they been long enough under uniform environments, to have had opportunity to 'notice' uniformly items collectible as tests. In respect to this third point, *school children* are a different group from *preschool children,* since the former will have been exposed to uniformities of environment (in the ordinary sense) that make possible a reservoir of items from which to construct tests.

It is to be noted, in criticism of the work of Wellman, that this investigator does not discuss these limitations of tests made at Iowa on preschool children. She does not take into consideration that tests standardized for preschool children are always brought out on less than a random sample. She uncritically accepts tests developed years ago for preschool ages as being tests standardized on the same samples of population as those now used during school years, which is not the case. Again, Wellman never considers the influence of negativism (resistance) at all, although Rust showed, in 1931, that when three-year-olds are being considered, 69 percent, by Kuhlmann-Binet, and 30 percent, by Merrill-Palmer, show significant reductions of IQ on account of negativism alone. Nor does she refer to the work of Reynolds and of Rust, showing how large is the factor of negativism with preschool children. The effect of overcoming resistance in raising the IQ is graphically presented. Nor does the age of this group define the peak of negativism. Younger children are even more resistant.

In practice these various sources of error work out, of course, as *variable* influences. The work of Jones shows this particularly well. Some IQ's will rise; some will fall; some will remain constant. It is hard to tell why Wellman's results differ so materially from those of other's who have worked with preschool children, in finding such large differences. On the whole, IQ's taken at preschool ages would be expected to *rise* for children in nursery school and after as a group, because (1) the preschool tests have a too high standard, and (2) the negativism of the children would tend to diminish through familiarity with adults in the school and with the school situation. And in accordance with this expectation, other investigators do get a slight rise in IQ in nursery-school children when compared with control groups, though not nearly to the extent found by Wellman.

It seems probable, as Simpson, McNemar, and Terman, among others, have pointed out, that the different results of Wellman are due to uncritical neglect of the *fallacy of selection.*

Since Wellman's results in modifying IQ's are so striking, and since they would have such great importance for social science and for education if verified, it becomes incumbent upon her to make clear to her professional colleagues what factors in the situation at Iowa produce the modifications. This clarification would presumably take the form of a discussion of 'environment.'

Wellman does, in fact, undertake a discussion of environment in Chapter II of Part I. Her discussion contains much that is stimulating, and it shows sustained thought on the concept of which it treats. The definition of 'stimulating environment' must, however, be uncompromisingly rejected, as a case of reasoning in a circle.

According to Wellman, a stimulating environment is one that stimulates; that is, one that succeeds in raising the intelligence quotient. This leaves us with no possibility of obtaining negative results for the nurture problem. If we fail to raise intelligence levels, this must be due to the fact that our environment is unstimulating, since by *definition* a stimulating environment is one that stimulates.

If we were to accept this point of view, we should be left with queer problems on our hands. Persons deteriorating from paresis or old age would *ipso facto never* be in a stimulating environment. Persons of mature intelligence would be forever immune from stimulating environments, since the intelligence level of such persons no longer rises.

This is like offering a definition of fire as "that which burns things." But fire does not burn a rock or a sheet of asbestos. Is it not, therefore, fire? Yes; it is fire just the same, for fire burns only combustible matter. Fire does not cease to be fire, just because the rock will not burn.

Furthermore, an environment does not cease to be, or fail to be, stimulating just because an IQ fails to rise. It may be that the IQ is not modifiable, just as the rock is not combustible.

I therefore reject Wellman's definition of 'stimulating environment' as circular.

These are the main items to be mentioned in a limited appraisal. Especially valuable features of the Yearbook are the clear definition of 'intelligence,' by Dr. Freeman, and certain of the more extended and carefully planned studies in Part II.

Crucial studies of the rôles of nature and of nurture take so much money, time, and critical insight that none has yet been made, nor can we expect them in the near future. Enough money, enough time, and enough insight for this task will be hard to secure and to bring together in such a way that the task will be carried out. Eventually, we have faith to believe, this essential combination is bound to occur.

V

Harold E. Jones

To those who are seeking final answers to questions about mental growth, the present Yearbook can give little satisfaction. The evidence will seem to them ambiguous and insecure; the clash of opposing points of view will suggest an even more fundamental insecurity in basic concepts. Does it appear likely that the present difficulties will be resolved in a Yearbook to appear in some later decade? Or should we, perhaps, advise potential contributors that this has now been shown to be an unproductive field of research?

It is the writer's view that we should not become impatient with research achievements, so long as effective gains continue to be made in techniques of measurement, in methods of inquiry, and in theory.

With regard to the first of these, some encouragement can be derived from the fact that our instruments for measuring intelligence show certain definite advances, and our use of these instruments has become more discriminating. We are less naïve than formerly as to the meanings of total scores, or as to the comparability of different scales at the same age level or of the same scale at different age levels.

With regard to our research methods, encouragement can also be derived from the fact that we are moving in the direction of a more effective experimental design. The studies of foster children and of separated identical twins have involved experimental comparisons on an *ex post facto* basis, without actual control over the experimental factors. Research on the influence of the nursery school has for the most part been on a similar basis, but has opened opportunities for further work in which specific experimental issues can be formulated and the pattern of variables can to some extent be controlled. We have passed beyond the stage in which it seemed feasible to determine in any *general* way the relative contributions of heredity and environment to variations in intelligence, for it has become only too evident that these relative contributions vary widely according to the characteristics of the test and also of the sample (age level, hereditary background, environmental background). In view of these complexities it seems probable that we shall turn from retrospective surveys of conditions assumed to have had a prior influence, and shall prefer to deal with the current and cumulative effects of specific experimental factors. It may also be expected that our interest will shift to some extent from mass statistical studies or purely descriptive accounts of mental growth curves in individuals to investigations of the dynamics of the growth process in individuals. We are now clear as to the common occurrence of fluctuations, upward trends, and downward trends in the seriatim mental scores obtained for selected cases. We are not at all clear as to the sources of these fluctuations and trends, much less as to means of controlling them; a frontal attack is needed in this field.

The development of theory has not as yet given a commanding position to the doctrine of a 'general' factor, of 'specific' factors, or of 'group' factors. The present volume is for the most part concerned with results based upon total scores from standard, generally accepted intelligence tests; little consideration is given to comparative studies of different types of test items or to analytic studies that attempt to deal with independent factors such as 'g,' 'V,' and so forth. From the point of view of practical implications, one cannot doubt the importance of working with composites that have already been shown to represent educationally significant samplings of abilities. It seems, however, probable that fundamental nature-nurture inquiries will in the future include studies based on tests more homogeneous as to content than the tests now in most common use.

To summarize, it is the view of this writer that the research of the past twelve years, since the publication of the 1928 yearbook of this Society, has made sufficiently promising advances to justify continued active work in this field; it is predicted that further work will be in the direction of (*a*) group studies involving more rigorously controlled experimentation, (*b*) studies of the dynamics of individual mental growth; that is, the processes directly influencing development, and (*c*) studies that will utilize measures of mental ability in terms of individual items or sub-tests or in terms of specific components.

VI

George D. Stoddard

Dr. Carmichael has observed that some differences in data, findings, and interpretations arise out of different methods of defining and analyzing the areas under investigation. Given exactly the same data and the same procedures, scientific workers should arrive at consistent and intelligible formulations, but in controversial fields even this is difficult to achieve.

The comments that follow are in no sense a full criticism of the contents of the Yearbook; they are offered simply as illustrations of semantic variations upon the themes of eight of the papers.

1. The Goodenough and Maurer Paper on Mental Development of Nursery-School Children Compared with That of Children without Nursery-School Training

The Goodenough and Maurer article presents an interesting sampling problem. In Table I data are presented for 147 nursery-school children and 260 non-nursery-school children. For various reasons, 134 of these nursery-school children and 245 of the non-nursery-school children were eliminated during the course of the study. For three years' attendance in a nursery school, this leaves a comparison between 13 nursery-school children and 15 non-nursery-school children. Even this small group is divided into three portions on the basis of occupational classification of the father.

All comparisons of intelligence quotients in this study are based on the Minnesota Preschool Scale, a test evidently not directly comparable to the Kuhlmann or Stanford revisions.

2. The Lamson Follow-Up Study of Nursery-School Children

The Lamson study shows a series of insignificant differences as between nursery school and non-nursery school, based on a small sampling. Since the non-nursery group was 2.7 months older than the nursery group, this factor alone might be expected to make a difference of 4 points in IQ for these selected children. Since Lamson's data indicate a difference of 3 points favoring the nursery-school group, what we possibly have is a 7-point advantage.

3. Hildreth on Adopted Children in a Private School

Hildreth's study includes children who had been poor learners and problem children long before they entered the Lincoln School, which for some was a ' last resort.' The sampling, if analyzed, might well include a mixture of genetic deficiency, birth injury, constitutional inferiority, and family rejection of a poorly developing child.

4. The Reymert and Hinton Study of the IQ's of 100 Children

In the Reymert and Hinton study of children, it is not clear that the conditions at Mooseheart are really superior for children, in terms of the usual dynamic and affectional relationships between children and adults. They state, for example, that there may be as many as fourteen young boys in a cottage under the care of a housemother and a cook. The fact that the children did not lose in IQ with residence in Mooseheart may be a point in favor of the institution, for there is evidence that the children studied might have *decreased* in IQ had they remained in their own homes; the children admitted at ages 10 to 14 years were mentally inferior to those admitted at age 6 or below.

Table III in the Reymert-Hinton article is noteworthy as one of the few in the research literature that fails to show a relationship between the IQ's of children and the economic status of the fathers. The IQ's of these Mooseheart children fell in the middle 90's for the three income samplings: $0 to $75 per month, $76 to $150 per month, and $151 to $225 per month. But the group of children who received regular schooling showed, on entrance to Mooseheart and throughout the period of this study, a superiority of 12 IQ points over those who received irregular schooling.

5. Stippich's Study of the Mental Development of Children of Feebleminded Mothers

It should be noted that this study is not directly comparable with the studies of Skeels and Skodak. Stippich does not present data indicating the mental growth of children of feebleminded mothers placed in boarding homes of superior occupational level. If these children, too, should turn out to be below average (as does her whole experimental group), the findings will be considered somewhat inconsistent with the Iowa studies. But the two sets of studies would not be contradictory, for the Iowa work is based entirely on the placement of children *for adoption in selected foster homes.*

In essence, Stippich reports in this preliminary study that, under equal conditions of stimulation, in institutions and commercial boarding homes, the children of normal mothers will test about average, while those of feebleminded mothers will be dull-normal (central tendencies). For the illegitimate children of normal mothers it will be helpful to have a further checking of the sampling and of the conditions incident to financial support, guardianship, selection of boarding home, and school facilities.

If, in the full sampling to be made available in a further study by Stippich, these results are borne out, it is possible that both Stippich and Skodak will be right. In short, under certain conditions the children of feebleminded mothers may achieve normality (Skodak); and under other conditions, subnormality (Stippich).

6. The Terman and Oden Study of the Correlates of Adult Achievement in the California Gifted Group

Dr. Terman in his follow-up study differentiates an *A*, or successful group, from a *C*, or unsuccessful group. Concerning these he remarks: " favorable also from the eugenic point of view is the fact that the A's appear to marry better [than the C's]." This is confusing. If one believes in strong *genetic* determiners of mental ability, why is such marrying favorable? *The A's and the C's had the same mental ability;* if this was inherited, they can presumably transmit it to their offspring through ' good ' marrying.

Dr. Terman points out that from 30 to 40 percent of the parents of these gifted children report mental abnormality among near relatives. Since a questionnaire is likely to show minimal figures on such traits, one is led to ask; Are these bright children, who are themselves stable, appearing in family lines strongly marked with instability?

Dr. Terman's whole article could be interpreted as supporting the dominance of social over genetic heredity.

7. Hollingworth and Terman on the Significance of Deviates

I should agree with Hollingworth and Terman that " low intelligence causes both poor environment and delinquency "; but the low intelligence is shown chiefly on the part of citizens and civic authorities who design, maintain, or protect slums and other impoverishing conditions.

8. John Anderson on the Prediction of Terminal Intelligence

Dr. Anderson's chapter raises some interesting theoretical questions. He holds, for example, with respect to mental testing " that in making the best possible prediction of terminal status, we will also make the best measurement of present status, in so far as our concern is with potentiality rather than with achievement." One might ask, ' terminal ' at what age? Is it true that we are really trying to predict how fast the child may become man, and as man, sink into senility? Perhaps our concern is with relative achievement *at any age.* In this sense estimates of potentiality arise from successive measures rather than from long-range prediction. It would appear that achievement in the young child is as valid as achievement later on. Is not the learning of the first fifty words around the age of two as characteristic of intelligent activity as any subsequent learning?

H. V. Meredith has shown that highly reliable and valid measures of standing height do not furnish a good basis of prediction through the chronological ages 11 to 18 years. According to Anderson's thesis, this would throw doubt upon the validity or usefulness of the measures at age 11. Is such a position tenable? Following Anderson's analogy, what would *true height* or *true weight* signify?

To the extent that Anderson's position is defensible, revisers and makers of tests are delinquent in claiming validity for testing at any age before they have obtained a satisfactory correlation with subsequent ages. Of course, there is no good reason for stopping at age 18 or 19; one would have to ask, do these tests, say at age 18, show a good prediction of mental ability at age 50?

It should be recalled, too, that the inconstant IQ (which is a fact and not a postulate) may indicate not only the measurement of different functions at different chronological ages, but also actual changes in the rate of growth of any ability or cluster of abilities.

VII

Lewis M. Terman

As Chairman of the Committee of the Society that prepared the 1928 Yearbook on the nature-nurture issue, I am naturally interested in comparing the evidence as it stood in 1928 and as it stands today.

The intervening years have not been as fruitful of crucial research as one could have hoped for. The period has in fact been marked by instances of retrogression in the methodological procedures that have been employed. This has shown itself at times in careless formulation of the problems to be attacked, use of unsound techniques, neglect of proper controls, misinterpretation of data, and the publicizing of unwarranted conclusions. But there has been good work as well as shoddy work, and in some respects I believe that genuine progress has been made.

One who has followed the age-old controversy between hereditarians and environmentalists will hardly be surprised to find that the new evidence does not all point in one direction. Its inconsistency is partly accounted for by the difficulty of controlling all the variables that may affect the results of a given experiment. The rule seems to hold that the more carefully the irrelevant factors are controlled, the weaker the evidence for large environmental influences. In this respect the situation is strikingly like that in the field of E. S. P. experimentation. Thus Freeman's foster children study, which left the door wide open for selective placement, seemed to indicate rather large environmental effects. The contemporary study by Burks, which largely ruled out this factor by limiting subjects to those adopted under the age of twelve months, indicated very slight effects. It is of singular interest that the later study by Leahy, modeled after that of Burks, but even more rigidly controlled, agrees with the latter almost perfectly to the second decimal in the low correlations found between intelligence scores of foster children and their foster parents. Wellman, leaving uncontrolled a half dozen or more important variables, finds large IQ effects from school attendance; Goodenough, Jones, and others who control the irrelevant factors find none.

For a more detailed account of the nature-nurture evidence appearing between 1928 and 1940 I refer the reader to Goodenough's two chapters in Part I, which are characterized by critical insight, scientific caution, and just appraisal of others' work. In contrast, the chapter by Stoddard and Wellman (XIV, Part I) impresses me as biased and uncritical. Results in line with their views are consistently played up, while contradictory results are played down or ignored. ' Favorable ' data are taken at their face value when alternative explanations are obviously possible in terms of such factors as selected population, statistical regression, practice effects, subject rapport, examiner bias, unreliability and standardization errors in the intelligence scale used, shift from one scale to another in retests, or the doubtful validity of intelligence scales for infants.

A similar bias permeates Wellman's chapter on schooling effects (XXVI, Part II), and to a somewhat less extent her chapter on the nature of environment (II, Part I). It appears characteristic of the Iowa group of workers that they so often find difficulty in reporting accurately either the data of others or their own.

In view of the fact that the contents of this Yearbook have been so largely shaped by controversy over the Iowa investigations, an evaluation of the present status of our problem calls for frank criticism of the methods these investigators have used and of the conclusions to which they have led. The latter are sensational in character and have been widely publicized by the authors through both lay and professional channels. If the discoveries proclaimed are genuine, they are among the most significant of all time.

How genuine they are will be difficult for the reader to ascertain from even a careful examination of the original reports, since the raw data have in most cases not been published. Anyone who is seriously interested in the validity of the claims should read not only the original documents (so far as these can be read intelligently without access to the raw data), but also Goodenough's chapter on sources of error in such studies (XII, Part I), J. Anderson's timely and clarifying discussion on the limitation of infant tests (XIII, Part I), and McNemar's thoroughgoing critique of the Iowa studies (*Psychological Bulletin*). The reader cannot fail to be impressed by the number of variables left uncontrolled in these investigations, by the faulty statistics employed, and even more by the extraordinary discrepancies between data presented and the conclusions drawn. A few examples follow.

In the orphanage preschool study Wellman and her co-authors claim that on the one hand, " distinct differences in patterns of growth were found between preschool and control children "; that " the effect of long [orphanage] residence for the control children was a leveling one, tending to bring all children to high-grade feeblemindedness or border-line classification"; that on the other hand, preschool attendance "exerted a profound influence" upon children of the experimental group, some of whom "were made placeable in foster homes who would have been doomed to commitment to an institution for the feeble-minded. . . ."

Just how devastating the orphanage life was, or how profound the influence of preschool attendance, could not be learned from the 190 pages of tables and conclusions published, but when the original scores were made available upon request, it was found that the average change from initial to final test was a loss of only 1.3 IQ points for the control group and a gain of only 1.8 IQ points for the preschool group. The proportion of subjects who gained was almost the same for the control group as for the preschool group (46 percent as against 51 percent) and gains of 16 points or more were actually twice as frequent in the control group (14 percent as against 7 percent). The imaginary character of the alleged "leveling effect" of the orphanage environment was further shown by the fact that variability of the central group on the final test differed by less than one percent from the variability on initial test. In short, there is no reason to believe that preschool attendance or non-attendance by these children had any effect whatever on their mental growth patterns. A single table giving the original scores of the subjects would have told the story.

As another example, consider the claims by Skodak, Skeels, and Stoddard regarding the high IQ's produced in children of low-grade parents by early adoption into superior foster homes. The picture presented is that of children of parents close to borderline deficiency being quickly brought to the level of children belonging to the professional classes. This picture is by no means consistent with the facts given. Both the mothers of these illegitimate children and the fathers, so far as the fathers were known, had completed on the average 10 grades of schooling and had a mean occupational rating very little below the general population of corresponding age. The 80 mothers who had been given a 1916 Stanford-Binet test earned an average IQ of 88, which, for adult subjects, is equivalent to about 94 on the revised Stanford-Binet. If slight allowance be made for the fact that the tests were administered within a few days of the illegitimate birth, the conclusion suggested is that the tested mothers were of average ability.

Furthermore, it turns out that the IQ's of these foster children were not superior, but instead approximately average. The mean usually quoted by the Iowa investigators is 116 (sometimes 120), but when we omit the tests given under the age of 24 months, below which level the test used is poorly standardized and of negligible validity as a measure of intelligence, the mean becomes 108. This is only one point above the mean of Kuhlmann's standardization group and three points above the mean of Goodenough's large unselected group. The true IQ mean of the group was not 16 points, but probably in the neighborhood of one to three points, above the generality. A mean at this level is not surprising, since Leahy has shown that illegitimate children offered for adoption average higher in mental ability than illegitimate children in general.

Another finding emphasized in Skodak's study is that the children adopted into the best foster homes averaged 12 points higher in IQ than those in the poorest homes. A large part of this difference was clearly due to selective placement. There was a correlation of .33 between child's IQ and education of the *true* parents, which is as high as Burks and Goodenough found for children living in their own homes. There was a correlation of .35 between occupation of true parent and of foster parent.

The most challenging of Skodak's conclusions is that " children of definitely moronic mothers, if placed early in good foster homes, will turn out to be above average in mental ability." This sweeping statement is based upon 16 children whose mothers had been classified as feebleminded. Whether the mothers were in fact feebleminded cannot be judged by the evidence reported. In view of all that is known regarding the inheritance of mental deficiency, so broad a generalization from a handful of cases is hardly to be taken seriously.

Confidence in the Iowa studies is further weakened by the numerous instances in which the authors have utilized unsound statistical procedures. Examples pointed out by McNemar include, among others, the presentation of selected data, the use of percentile scores for correlational purposes, the averaging of percentile scores, the use of inflated N's, the misinterpretation of regression phenomena, the mixing of doubtfully comparable test scores, and the total neglect of evidence pointing to the limited validity and low reliability of infant tests.

One is especially puzzled by the persistence with which these investigators have taken for granted the equivalence of test scores in infancy and later childhood. Bayley's retest data published more than six years ago showed that scores obtained before 18 months are almost totally uncorrelated with those obtained at ages beyond four years, and that only after the age of 24 months do the scores begin to acquire appreciable predictive value. Bayley's results have been amply confirmed by others. The truth is that an infant destined to develop into a moron may test as high at 18 months as the infant who will ultimately make the 140 IQ class. The Iowa investigators have capitalized heavily on the low validity and low reliability of infant and preschool test scores.

Skeels and Dye have carried this technique to absurd limits in their claim that feebleminded children (diagnosed as such at an average of 17 months) were quickly made normal as a result of treatment administered by moron nursemaids. We are informed that this novel system of tutelage raised the mean IQ from 64 to 92. From the point of view of regression phenomena, it is interesting to note that practically all the gain (26 of the 28 points) was accomplished in 9.4 months. We are further asked to believe that, but for this special treatment, the group would have *decreased* in IQ by some 26 points, which would mean a net accomplishment of 54 IQ points saved and gained as a result of the moron tutelage ($28 + 26 = 54$). The claim of 26 IQ points saved is based on tests of what the authors call a "contrast group." This appears to be a new technique in the history of experimental psychology, for the "control group" was made up, not at the time the experiment started, but after it was over, and on the basis of test scores already on the books. Such an *ex post facto* selection of cases is indefensible in view of the fact that, as we have previously shown, the Wellman control group in this same orphanage did not show a reliable drop in IQ.

It happens that the Yearbook contains another contribution on the IQ effects of association with the feebleminded (XXI, Part II). There we find reported a mean IQ of 100 for children who had lived less than two years with their feebleminded mothers, and a mean of only 52 for those who had remained with their defective mothers to the age of 12 years. The difference is 47 points. If we add this 47-point drop caused by association with defective mothers to the Skeels-Dye 28-point increase caused by association with moron nursemaids, we get 75 IQ points as a measure of the difference in intellectual stimulus value between these two classes of defective females. This may not seem to make sense, but it is at least an interesting wonderland that the environmentalists have opened to us.

An alleged discovery in science stands or falls according to whether it can be confirmed with some regularity by other investigators. Judged by this criterion, the Iowa claims regarding the influence of schooling on mental development are subject to gravest doubts. Of nine new studies reported in Part II on the effects of nursery-school experience, eight are entirely negative and the ninth practically so. Starkweather and Roberts, using the unsound statistical procedures common to the Iowa studies, interpret their data as confirming Wellman's conclusions. However, the IQ increase of 4.1 points shown by their data could all have been caused by practice effects and decreased shyness of the subject. Several of the preschool studies, notably those of Goodenough and Maurer, Jones, and L. D. Anderson are models of their kind. In all the literature of child psychology it would be hard to find a more impressive contrast in scientific method than that offered by the Minnesota and Iowa investigations of nursery-school attendance.

The new studies dealing with school children above the nursery-school age are also rather consistently negative. Pritchard, Horan, and Hollingworth (XVI, Part II) report complete inability to lift the IQ's of dull children in special classes by the most stimulating program they were able to provide. Hollingworth is unable to increase the learning ability of gifted children at 130 IQ to the learning ability at 160 IQ under a special class program of instruction that probably has no equal in this country. Hildreth (X, Part II) finds that the Lincoln School of Teachers College, conducted with all the expertness that scientific pedagogy can bring to bear, is unable to lift above average the IQ's of foster children whom its wealthy and cultured patrons have adopted and reared as their own. Thorndike's report (XXIV, Part II) of retests given to 1,166 pupils in three leading private schools gives no real support to the Stoddard-Wellman point of view; the one school that showed a significant IQ gain (6 points) was the one whose examiners were least well trained and most numerous, and the one that Thorndike says could not be regarded as the most intellectually stimulating.

Kephart's report (XIII, Part II) of a ten-point increase in the IQ's of feebleminded subjects who had been given special training is impossible to evaluate without more information than the author has given regarding his training program and the extent to which its content may have resembled that of the Binet tests.

There are two possible explanations of the discrepant data reported by Wellman: either the Iowa educational program is more stimulating than that of other schools or her findings are invalid. On the first alternative, a survey by an expert educational commission would be helpful. In the absence of such expert opinion one is thrown back on the published descriptions of the curricula and methods of the schools in question. The accounts I have been able to find do not suggest that the Iowa educational procedures are particularly unique or superior to those employed in other schools where the problem has been investigated with negative results. Unless this judgment can be shown to be incorrect, we are left with the alternative explanation.

On the whole, I think that the evidence accumulated during the last dozen years has weakened rather than strengthened the position of the environmentalist. I base this opinion not only on the positive evidence for hereditary influences but even more on the growing recognition of weakness and error in the evidence that environmentalists have most heavily stressed. We are certainly far short of having adequate proof that the rate and character of mental maturation are determined primarily by exogenous rather than endogenous factors.

This is not to say that environment can be regarded as a matter of small consequence. To it alone we owe the major cultural advances that separate us from our barbaric ancestors of 3000 years ago. The hereditarian can wholeheartedly join with the environmentalist in the demand that more and better educational opportunity be provided for every child, but he has too much respect for facts and too little faith in miracles to expect that equalization of opportunity will result in equality of achievement.

It is unfortunate that the controversy should have become so exclusively concerned with environmental influences upon the IQ. An obtained IQ is not only subject to chance errors resulting from inadequate sampling of abilities, but also to numerous constant errors, including practice effects, negativism, or shyness, the personal equation of the examiner, and standardization errors in the test used. For these reasons an obtained IQ, as I have many times pointed out, should never be taken as a final verdict, but only as a point of departure for further investigation of a subject. The great mistake of the Iowa workers is that in practice they use obtained IQ's as though they were always true measures of present ability, regardless of the subject's age and regardless of the limited reliability or validity of the scale employed. Any IQ changes observed are accepted as representing true changes in rate of mental growth and are invariably explained in terms of 'environmental impacts.'

It is easy to imagine deprivation so great that an IQ score would have no meaning whatever. Such would be the case with a child reared in a cage, in complete isolation, by some kind of robot that provided the necessary care to keep him alive and healthy. It is another question whether a child so deprived during his early years would be permanently robbed of his potentialities for normal mental growth. For all we know, he might quickly realize his genetic potentialities if moved into a favorable environment. An experimental approach to this problem is offered by identical twins whose environments have differed so radically as to be reflected in their IQ's. It would be extremely interesting to know whether IQ differences so caused would disappear after a period of equalized educational opportunity. It is to be hoped that someone will undertake an experiment of this kind. My prediction is that when the experiment is made with subjects not too old (under 15, say) the IQ differences will be reduced to something like the magnitude commonly found for identical twins reared together.

The issue is not simply whether IQ's can be influenced by differences in environment and training. That to some degree they are so influenced, no one has ever denied. Whether in a typical American community the influence is relatively small (as I believe) or quite large (as some believe) is less important than whether it has a permanent effect upon capacity for achievement. On this question the evidence from life histories is most crucial. Biography affords many examples of notable or even prodigious achievement by men and women who in childhood suffered extreme educational deprivation. Helen Keller, completely isolated from the world of ideas until well beyond the preschool period, was nevertheless able to graduate from college with highest honors and to become the author of books and poems of distinguished literary quality. Anne Sullivan at 14 was blind, entirely illiterate, and had lived for several years in one of the worst orphanages this country has known. Removed to the Perkins Institute at 14 and given her sight a few years later by a surgical operation, she was selected at 22 for one of the most difficult educational tasks ever undertaken. The life and works of Helen Keller are the monument to her pedagogical genius. Case histories of this kind render almost foolish the belief that intellectual potentialities are permanently affected by a little more or a little less attendance at a particular nursery school.

VIII

Beth L. Wellman

The reader who has perused the chapter on new evidence on evironmental influence prepared by Goodenough and the chapter on the same topic prepared by Stoddard and Wellman will realize that there is a wide difference of opinion. The reader need not be disturbed by this; he may be encouraged to draw his own conclusions from the evidence.

The Iowa point of view has been modified gradually by new lines of evidence bearing on environmental conditions. It may be appropriate to point out that most of the Iowa results were unexpected; they were not what had been anticipated. In some cases the results were disconcerting, as, for example, when Skeels found that two children, diagnosed as mentally retarded and committed to an institution for the feebleminded with the expectation that the commitment would be permanent, increased markedly in IQ. But the findings on these two children led directly to a further study of a larger group. This confirmed the results and gave new information on the dynamics of environmental factors. It was embarrassing, too, in another investigation to report to the superintendent of an orphanage, who had graciously granted full research privileges, that his institution was unfavorable to child development at the younger ages.

An independent duplication of the Iowa studies would be welcomed, with every effort made to determine the true causes of change. It is suggested, too, that a similar check be made on situations where investigators report no change. Why, in any sampling, including those that show little mass upward or downward shifting, should there be individual tendencies toward acceleration or deceleration? Perhaps one of the main contributions of the Yearbook is to point out the complexity of the problem and the inadequacies of our measuring instruments, both of intelligence and of environmental factors.

Concerning Goodenough's discussion of sampling errors, I should like to raise the question: What effect do sampling errors have upon the child's *change* in IQ from one measurement time to another? Possibly Goodenough has confused level of development with change. It is well known that children from the higher occupational classes have a higher mean IQ than those from the lower classes. But when studies begin with groups that are equal in initial IQ, the question becomes:

Will differences in the quality of domicile and in the economic security of parents be related to *change* in IQ (rate of subsequent growth), so that two groups of children initially equal (presumably equal in previous growth) will now grow at different rates? Does it not follow that, in her insistence that home backgrounds are related to differences in growth potentialities, Goodenough has placed herself in a position against the constancy of the IQ?

Since a large portion of Goodenough's chapter is devoted to a discussion of the Iowa studies, perhaps some specific comments are in order here. One contribution to Part II by Wellman contains the summary of the Iowa studies on the effects of schooling, and another by Skeels is a summary of studies on foster children. The serious student of the question is asked not to depend on these brief summaries, but to refer to the original researches from which they were taken.

In regard to the Iowa preschool studies, it should be pointed out that children have been studied in a control situation; namely, that of non-attendance (see my summary in Part II, Chapter XXVI). Comparisons have been made of the change in IQ of the same children under different conditions (attendance and non-attendance). It is felt that this type of comparison offers controls as adequate as comparison between different groups. Certainly it reduces the problems inherent in sampling errors, selection of children, inadequacies of matching, and differences in home background, since all of these are taken care of by the fact that the children are identical. However, these results have been supplemented by comparisons of control groups of non-preschool children.

The validity of changes at the preschool ages cannot be determined entirely by the permanence of any IQ ratings attained. Later development, too, is dependent to some extent on the kind of environmental conditions that obtain. Some groups of children who have gained in IQ during preschool make further gains in the elementary-school years. Some keep the same IQ. Conceivably, still other groups would lose in IQ if they lived under unfavorable conditions. Our studies lead to the conclusion that, so far as is now known, under ordinary circumstances, the IQ accelerated through preschool attendance is maintained.

Goodenough has suggested that practice effects may account for the gains made during the elementary-school years. May I call attention to the fact that generously scattered through the Iowa studies are analyses in which the number of tests has been held constant, and that number of tests does not explain differences in rate of growth?

The problem of regression has been discussed by Stoddard and Wellman in Part I, Chapter XIV. If regression were the sole explanation of differential changes, two groups of children with the same initial test scores should show the same changes. As a matter of fact, they do not, when environmental conditions are different. The changes are not the same, for example, for preschool children from fall to spring (period of attendance) as from spring to fall (period of non-attendance). But there is a need for tests that are more reliable, since regression effects emerge from test unreliability.

I wish to make two or three points by way of clarification.

The reader will find that the gain made in four years' attendance at the University of Iowa Elementary School was 5.6 points, and that this amount of gain was made by a group already very superior.

In regard to the IQ changes in orphanage preschool children, the authors of the study showed that the preschool and non-preschool groups were not differentiated by short periods of residence, but that there were marked differences over longer periods. Over the longest residence period, three preschool children and one non-preschool child gained 15 or more points. The greatest losses in the preschool group were by three children who lost 13, 14, and 15 points. The greatest losses in the control group were by seven children who lost 11, 14, 14, 15, 23, 37, and 43 points. When tests under twenty months of age are omitted, the loss by the non-preschool group, initially 80 or above in IQ, who were in residence 400 days or more remains statistically significant, as does the difference between the change in the non-preschool and the preschool groups.

That children of average ability in the orphanage preschool study descended to subnormal levels is not a theory, but a simple fact. For example, one child initially average lost 43 points in IQ; another, 37.

Perhaps a word should be said about the Peterson [1] study. Children entering the junior primary group (five-year-olds) in 1935 were divided into two groups, those who had attended preschool and those who had not. The two groups were similar in IQ on the fall test in

[1] T. J. Peterson. "A preliminary study of the effects of previous nursery-school attendance upon five-year-old children entering kindergartens." [In] R. Updegraff, M. E. Keister, L. Heiliger, and Others. *Studies in Preschool Education, I.* (University of Iowa Studies in Child Welfare, Vol. 13, No. 4; 1937, pp. 197–248.)

junior primary. The fact that the two groups did not differ greatly in reading readiness, vocabulary, and information is not surprising, since these are related to IQ. Even though the two groups did not differ significantly on entering junior primary, the preschool group had already gained in IQ. The previous history of IQ changes of the non-preschool group was unknown.

INFORMATION CONCERNING THE NATIONAL SOCIETY FOR THE STUDY OF EDUCATION

1. Purpose. The purpose of the National Society is to promote the investigation and discussion of educational questions. To this end it holds an annual meeting and publishes a series of yearbooks.

2. Eligibility to Membership. Any person who is interested in receiving its publications may become a member by sending to the Secretary-Treasurer information concerning name, title, and address, and a check for $3.50 (see Item 5).

Membership is not transferable; it is limited to individuals, and may not be held by libraries, schools, or other institutions, either directly or indirectly.

3. Period of Membership. Applicants for membership may not date their entrance back of the current calendar year, and all memberships terminate automatically on December 31, unless the dues for the ensuing year are paid as indicated in Item 6.

4. Duties and Privileges of Members. Members pay dues of $2.50 annually, receive a cloth-bound copy of each publication, are entitled to vote, to participate in discussion, and (under certain conditions) to hold office. The names of members are printed in the yearbooks.

5. Entrance Fee. New members are required the first year to pay, in addition to the dues, an entrance fee of one dollar.

6. Payment of Dues. Statements of dues are rendered in October or November for the following calendar year. Any member so notified whose dues remain unpaid on January 1, thereby loses his membership and can be reinstated only by paying a reinstatement fee of fifty cents, levied to cover the actual clerical cost involved.

School warrants and vouchers from institutions must be accompanied by definite information concerning the name and address of the person for whom membership fee is being paid. Statements of dues are rendered on our own form only. The Secretary's office cannot undertake to fill out special invoice forms of any sort or to affix notary's affidavit to statements or receipts.

Cancelled checks serve as receipts. Members desiring an additional receipt must enclose a stamped and addressed envelope therefor.

7. Distribution of Yearbooks to Members. The yearbooks, ready prior to each February meeting, will be mailed from the office of the publishers, only to members whose dues for that year have been paid. Members who desire yearbooks prior to the current year must purchase them directly from the publishers (see Item 8).

8. Commercial Sales. The distribution of all yearbooks prior to the current year, and also of those of the current year not regularly mailed to members in exchange for their dues, is in the hands of the publishers, not of the Secretary. For such commercial sales, communicate directly with the Public School Publishing Company, Bloomington, Illinois, which will gladly send a price list covering all the publications of this Society and of its predecessor, the National Herbart Society. This list is also printed in the yearbook.

9. Yearbooks. The yearbooks are issued about one month before the February meeting. They comprise from 600 to 800 pages annually. Unusual effort has been made to make them, on the one hand, of immediate practical value, and on the other hand, representative of sound scholarship and scientific investigation. Many of them are the fruit of coöperative work by committees of the Society.

10. Meetings. The annual meeting, at which the yearbooks are discussed, is held in February at the same time and place as the meeting of the American Association of School Administrators.

Applications for membership will be handled promptly at any time on receipt of name and address, together with check for $3.50 (or $3.00 for reinstatement). Generally speaking, applications entitle the new member to the yearbook slated for discussion during the calendar year the application is made, but those received in December are regarded as pertaining to the next calendar year.

Box 822, Clifton, Mass. GUY M. WHIPPLE, Secretary-Treasurer.

PUBLICATIONS OF THE NATIONAL HERBART SOCIETY

(Now the National Society for the Study of Education)

	Postpaid Price
First Yearbook, 1895	$0.79
First Supplement to First Yearbook	.28
Second Supplement to First Yearbook	.27
Second Yearbook, 1896	.85
Supplement to Second Yearbook	.27
Third Yearbook, 1897	.85
Ethical Principles Underlying Education. John Dewey. Reprinted from Third Yearbook	.27
Supplement to Third Yearbook	.27
Fourth Yearbook, 1898	.79
Supplement to Fourth Yearbook	.28
Fifth Yearbook, 1899	.79
Supplement to Fifth Yearbook	.54

PUBLICATIONS OF THE NATIONAL SOCIETY FOR THE STUDY OF EDUCATION

	Postpaid Price
First Yearbook, 1902, Part I—*Some Principles in the Teaching of History.* Lucy M. Salmon	$0.54
First Yearbook, 1902, Part II—*The Progress of Geography in the Schools.* W. M. Davis and H. M. Wilson	.53
Second Yearbook, 1903, Part I—*The Course of Study in History in the Common School.* Isabel Lawrence, C. A. McMurry, Frank McMurry, E. C. Page, and E. J. Rice	.53
Second Yearbook, 1903, Part II—*The Relation of Theory to Practice in Education.* M. J. Holmes, J. A. Keith, and Levi Seeley	.53
Third Yearbook, 1904, Part I—*The Relation of Theory to Practice in the Education of Teachers.* John Dewey, Sarah C. Brooks, F. M. McMurry, *et al*	.53
Third Yearbook, 1904, Part II—*Nature Study.* W. S. Jackman	.85
Fourth Yearbook, 1905, Part I—*The Education and Training of Secondary Teachers.* E. C. Elliott, E. G. Dexter, M. J. Holmes, *et al*	.85
Fourth Yearbook, 1905, Part II—*The Place of Vocational Subjects in the High-School Curriculum.* J. S. Brown, G. B. Morrison, and Ellen H. Richards	.53
Fifth Yearbook, 1906, Part I—*On the Teaching of English in Elementary and High Schools.* G. P. Brown and Emerson Davis	.53
Fifth Yearbook, 1906, Part II—*The Certification of Teachers.* E. P. Cubberley	.64
Sixth Yearbook, 1907, Part I—*Vocational Studies for College Entrance.* C. A. Herrick, H. W. Holmes, T. deLaguna, V. Prettyman, and W. J. S. Bryan	.70
Sixth Yearbook, 1907, Part II—*The Kindergarten and Its Relation to Elementary Education.* Ada Van Stone Harris, E. A. Kirkpatrick, Maria Kraus-Boelté, Patty S. Hill, Harriette M. Mills, and Nina Vandewalker	.70
Seventh Yearbook, 1908, Part I—*The Relation of Superintendents and Principals to the Training and Professional Improvement of Their Teachers.* Charles D. Lowry	.78
Seventh Yearbook, 1908, Part II—*The Co-ordination of the Kindergarten and the Elementary School.* B. J. Gregory, Jennie B. Merrill, Bertha Payne, and Margaret Giddings	.78
Eighth Yearbook, 1909, Parts I and II—*Education with Reference to Sex.* C. R. Henderson and Helen C. Putnam. Both parts	1.60
Ninth Yearbook, 1910, Part I—*Health and Education.* T. D. Wood	.85
Ninth Yearbook, 1910, Part II—*The Nurse in Education.* T. D. Wood, *et al.*	.78
Tenth Yearbook, 1911, Part I—*The City School as a Community Center.* H. C. Leipziger, Sarah E. Hyre, R. D. Warden, C. Ward Crampton, E. W. Stitt, E. J. Ward, Mrs. E. C. Grice, and C. A. Perry	.78
Tenth Yearbook, 1911, Part II—*The Rural School as a Community Center.* B. H. Crocheron, Jessie Field, F. W. Howe, E. C. Bishop, A. B. Graham, O. J. Kern, M. T. Scudder, and B. M. Davis	.79
Price for Yearbooks VI to X inclusive, 8 vo. cloth	5.00
Eleventh Yearbook, 1912, Part I—*Industrial Education: Typical Experiments Described and Interpreted.* J. F. Barker, M. Bloomfield, B. W. Johnson, P. Johnson, L. M. Leavitt, G. A. Mirick, M. W. Murray, C. F. Perry, A. L. Safford, and H. B. Wilson	.85
Eleventh Yearbook, 1912, Part II—*Agricultural Education in Secondary Schools.* A. C. Monahan, R. W. Stimson, D. J. Crosby, W. H. French, H. F. Button, F. R. Crane, W. R. Hart, and G. F. Warren	.85
Twelfth Yearbook, 1913, Part I—*The Supervision of City Schools.* Franklin Bobbitt, J. W. Hall, and J. D. Wolcott	.85
Twelfth Yearbook, 1913, Part II—*The Supervision of Rural Schools.* A. C. Monahan, L. J. Hanifan, J. E. Warren, Wallace Lund, U. J. Hoffman, A. S. Cook, E. M. Rapp, Jackson Davis, and J. D. Wolcott	.85

PUBLICATIONS OF THE SOCIETY

Thirteenth Yearbook, 1914, Part I—*Some Aspects of High-School Instruction and Administration.* H. C. Morrison, E. R. Breslich, W. A. Jessup, and L. D. Coffman.. .85

Thirteenth Yearbook, 1914, Part II—*Plans for Organizing School Surveys, with a Summary of Typical School Surveys.* Charles H. Judd and Henry L. Smith......... .79

Fourteenth Yearbook, 1915, Part I—*Minimum Essentials in Elementary School Subjects—Standards and Current Practices.* H. B. Wilson, H. W. Holmes, F. E. Thompson, R. G. Jones, S. A. Courtis, W. S. Gray, F. N. Freeman, H. C. Pryor, J. F. Hosic, W. A. Jessup, and W. C. Bagley.................................. .85

Fourteenth Yearbook, 1915, Part II—*Methods for Measuring Teachers' Efficiency.* Arthur C. Boyce.. .79

Fifteenth Yearbook, 1916, Part I—*Standards and Tests for the Measurement of the Efficiency of Schools and School Systems.* G. D. Strayer, Bird T. Baldwin, B. R. Buckingham, F. W. Ballou, D. C. Bliss, H. G. Childs, S. A. Courtis, E. P. Cubberley, C. H. Judd, George Melcher, E. E. Oberholtzer, J. B. Sears, Daniel Starch, M. R. Trabue, and G. M. Whipple... .85

Fifteenth Yearbook, 1916, Part II—*The Relationship between Persistence in School and Home Conditions.* Charles E. Holley...................................... .87

Fifteenth Yearbook, 1916, Part III—*The Junior High School.* Aubrey A. Douglass.. .85

Sixteenth Yearbook, 1917, Part I—*Second Report of the Committee on Minimum Essentials in Elementary-School Subjects.* W. C. Bagley, W. W. Charters, F. N. Freeman, W. S. Gray, Ernest Horn, J. H. Hoskinson, W. S. Monroe, C. F. Munson, H. C. Pryor, L. W. Rapeer, G. M. Wilson, and H. B. Wilson...................... 1.00

Sixteenth Yearbook, 1917, Part II—*The Efficiency of College Students as Conditioned by Age at Entrance and Size of High School.* B. F. Pittenger.................. .85

Seventeenth Yearbook, 1918, Part I—*Third Report of the Committee on Economy of Time in Education.* W. C. Bagley, B. B. Bassett, M. E. Branom, Alice Camerer, J. E. Dealey, C. A. Ellwood, E. B. Greene, A. B. Hart, J. F. Hosic, E. T. Housh, W. H. Mace, L. R. Marston, H. C. McKown, H. E. Mitchell, W. C. Reavis, D. Snedden, and H. B. Wilson... .85

Seventeenth Yearbook, 1918, Part II—*The Measurement of Educational Products.* E. J. Ashbaugh, W. A. Averill, L. P. Ayers, F. W. Ballou, Edna Bryner, B. R. Buckingham, S. A. Courtis, M. E. Haggerty, C. H. Judd, George Melcher, W. S. Monroe, E. A. Nifenecker, and E. L. Thorndike............................ 1.00

Eighteenth Yearbook, 1919, Part I—*The Professional Preparation of High-School Teachers.* G. N. Cade, S. S. Colvin, Charles Fordyce, H. H. Foster, T. W. Gosling, W. S. Gray, L. V. Koos, A. R. Mead, H. L. Miller, F. C. Whitcomb, and Clifford Woody 1.65

Eighteenth Yearbook, 1919, Part II—*Fourth Report of Committee on Economy of Time in Education.* F. C. Ayer, F. N. Freeman, W. S. Grey, Ernest Horn, W. S. Monroe, and C. E. Seashore.. 1.10

Nineteenth Yearbook, 1920, Part I—*New Materials of Instruction.* Prepared by the Society's Committee on Materials of Instruction............................... 1.10

Nineteenth Yearbook, 1920, Part II—*Classroom Problems in the Education of Gifted Children.* T. S. Henry.. 1.00

Twentieth Yearbook, 1921, Part I—*New Materials of Instruction.* Second Report by the Society's Committee 1.30

Twentieth Yearbook, 1921, Part II—*Report of the Society's Committee on Silent Reading.* M. A. Burgess, S. A. Courtis, C. E. Germane, W. S. Gray, H. A. Greene, Regina R. Heller, J. H. Hoover, J. A. O'Brien, J. L. Packer, Daniel Starch, W. W. Theisen, G. A. Yoakam, and representatives of other school systems............ 1.10

Twenty-First Yearbook, 1922, Parts I and II—*Intelligence Tests and Their Use.* Part I —*The Nature, History, and General Principles of Intelligence Testing.* E. L. Thorndike, S. S. Colvin, Harold Rugg, G. M. Whipple. Part II—*The Administrative Use of Intelligence Tests.* H. W. Holmes, W. K. Layton, Helen Davis, Agnes L. Rogers, Rudolf Pintner, M. R. Trabue, W. S. Miller, Bessie L. Gambrill, and others. The two parts are bound together... 1.60

Twenty-Second Yearbook, 1923, Part I—*English Composition: Its Aims, Methods, and Measurements.* Earl Hudelson.. 1.10

Twenty-Second Yearbook, 1923, Part II—*The Social Studies in the Elementary and Secondary School.* A. S. Barr, J. J. Coss, Henry Harap, R. W. Hatch, H. C. Hill, Ernest Horn, C. H. Judd, L. C. Marshall, F. M. McMurry, Earle Rugg, H. O. Rugg, Emma Schweppe, Mabel Snedaker, and C. W. Washburne.................... 1.50

Twenty-Third Yearbook, 1924, Part I—*The Education of Gifted Children.* Report of the Society's Committee. Guy M. Whipple, Chairman........................ 1.75

Twenty-Third Yearbook, 1924, Part II—*Vocational Guidance and Vocational Education for Industries.* A. H. Edgerton and others.................................. 1.75

Twenty-Fourth Yearbook, 1925, Part I—*Report of the National Committee on Reading.* W. S. Gray, Chairman, F. W. Ballou, Rose L. Hardy, Ernest Horn, Frances Jenkins, S. A. Leonard, Estaline Wilson, and Laura Zirbes.................... 1.50

Twenty-Fourth Yearbook, 1925, Part II—*Adapting the Schools to Individual Differences.* Report of the Society's Committee. Carleton W. Washburne, Chairman... 1.50

Twenty-Fifth Yearbook, 1926, Part I—*The Present Status of Safety Education.* Report of the Society's Committee. Guy M. Whipple, Chairman.......................... 1.75

Twenty-Fifth Yearbook, 1926, Part II—*Extra-Curricular Activities.* Report of the Society's Committee. Leonard V. Koos, Chairman.............................. 1.50

PUBLIC SCHOOL PUBLISHING COMPANY, BLOOMINGTON, ILLINOIS
Agents—The Baker and Taylor Company **New York**