EDUCATIONAL PSYCHOLOGY

EDUCATIONAL PSYCHOLOGY

AN INTRODUCTORY TEXT

BY

RUDOLF PINTNER

PROFESSOR OF EDUCATION, TEACHERS COLLEGE
COLUMBIA UNIVERSITY

Twinned edition.
New York: H. Holt & Co; 1929

LONDON
WILLIAMS & NORGATE LTD
1930

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

Educational Psychology is frequently considered an applied psychology, and strictly considered it undoubtedly is that. As an application of psychology to education, we may note that the earlier text-books of educational psychology followed rather slavishly the conventional order of subjects usually found in the text-books of pure psychology. Thus we find many early educational psychologies, and indeed several recent ones, beginning with a treatment of the nervous system, following on with a discussion of the several senses, then going on to such topics as perception, attention, emotion, memory, imagination and generally ending up with the higher thought processes. Frequently the applications to education in many of these topics were difficult to make, and often they were very far-fetched and had only the remotest connection with education.

Thorndike's *Educational Psychology* broke away very definitely from these traditions. His treatment was an investigation of original nature and its modifications. Thus we have the emphasis laid upon the original tendencies of the individual and the modifications of these original tendencies which the school is trying to bring about. This new departure of Thorndike's has influenced all educational psychologies written since the publication of his books. Some writers have taken over the new arrangement completely, others have tried a compromise between it and the older traditional sequence of topics in pure psychology, so that in some books we still see much space devoted to the nervous system and to sensation, perception and so forth. It is interesting to note here that this new departure of Thorndike's is now beginning to influence the text-books of pure psychology, and to break down the traditional treatment referred to.

Since the appearance of Thorndike's *Educational Psychology*, there has been a great development in the field of psychological and educational measurement. Tests of all kinds have been constructed. We are measuring to-day things that did not appear measurable fifteen or twenty years ago. This field of measurement has produced its own specialists and separate courses in measurement have sprung up in our colleges. To a great extent it has been kept apart from the regular course in educational psychology.

My belief is that this measurement of original tendencies and of educational modifications is an integral part of educational psychology. A study of original nature is helped by whatever measurements we can make, and a study of the learning process implies the measurement of educational achievement. It seems to me, therefore, that the time has come to incorporate such work into our educational psychologies, not as a separate topic added to the traditional treatment in a closing chapter, but rather as an integral part of the whole treatment. Some of the most important contributions of modern educational psychology are to be found in the recent field of measurement, and so it is only proper that an elementary text, such as this, should lay stress on these things. Many of our students do not have the time or inclination to take special courses in measurement in addition to the regular courses in educational psychology. It is my belief, therefore, that a first course in educational psychology should contain the general field of educational measurement as an integral part of the course. Throughout the book my aim has been to emphasize such topics as seem to me to be of most value to the teacher and educator. Some topics of doubtful value to the teacher, such as a study of the nervous system, have been omitted altogether. Curt and dogmatic treatment has been meted out to many controversial topics. Advanced students must go elsewhere, notably to Thorndike's books, for more extended treatment.

This book is designed for a one semester's course for students beginning a study of educational psychology. I have given a few advised readings at the end of most chapters. I have purposely kept these down to a minimum, with the hope that they will be read by all students. A lengthy bibliography at the end of each chapter is more likely to discourage rather than encourage further reading on the part of the beginning student. A general bibliography is given at the end of the book, which the teacher may use for additional reading assignments, if necessary. At the end of each chapter I have also given the references to specific pieces of work mentioned in the text. These references are not intended as reading assignments for all students, but merely as acknowledgments to the authors from whom I have borrowed and also as indications to the reader as to where he may go for original sources. I have also added a few exercises, mainly true-false statements, at the end of each chapter. These the student may use to practice on for review purposes. Some of the true-false statements will undoubtedly prove ambiguous and impossible to mark plus or minus, and such might readily stimulate discussion in the class.

As to acknowledgments, it seems needless for me to mention the debt that all educational psychology owes to Thorndike. The whole arrangement of this book and the treatment of most of the topics show his influence. I have made much use of pictures and diagrams borrowed freely from others, and for such borrowings I wish here to thank the following authors and publishers:

E. L. Thorndike, Leta S. Hollingworth, Daniel Starch, E. S. Robinson, Antoinette Feleky, Arthur I. Gates, Donald G. Paterson, Lewis M. Terman, W. F. Book, W. B. Pillsbury, H. E. Buchholz, Herbert Woodrow, W. H.

Pyle, Peter Sandiford, Harold E. Jones, Elizabeth Irwin, A. T. Poffenberger, E. M. Paulu, Henry H. Goddard, George D. Strayer, John A. McGeoch, Mary Wentworth McConaughy, Kathryn E. Maxfield, Carl E. Seashore, Walter Scott, Charles H. Judd, Arnold Gesell, Rachel Stutsman, Harry L. Hollingworth, William Healy, Bernard Muscio, Cyril Burt, C. H. Stoelting, The Macmillan Company, Stanford University Press, Psychological Review, D. Appleton and Co., World Book Company, Teachers College Bureau of Publications, Clark University Press, Silver Burdett and Co., Houghton Mifflin and Co., A. W. Shaw Company, National Research Council, Harvard University Press, Longmans Green, Warwick and York, D. C. Heath and Co., Cleveland Foundation, Archives of Psychology, National Society for the Study of Education, Journal of Applied Psychology, George Routledge and Sons.

RUDOLF PINTNER.

Yonkers, May, 1929.

CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

EDUCATIONAL PSYCHOLOGY

CHAPTER I

THE FIELD OF EDUCATIONAL PSYCHOLOGY

The Contribution of Psychology to Education. — The aim of all education from a psychological standpoint is to change the individual in certain directions. Whether we are thinking of the education of the child or the adult, we always are thinking of effecting some change in his reactions. Whatever be the instruments we may employ — school, church, newspaper, library, club — in every case we expect some change in the reactions of the individual to take place as a result of their use. In every case, then, our objective or goal is a changed reaction. In one case we may wish to teach a child to modify his reaction to the symbols 2×3 from saying or writing 5 or 4 or laughing or saying "I don't know" or turning away, to the reaction of saying or writing 6. In another case we may wish to modify the reaction of a voter on election day from using all the day for an outing, or thinking that it does not matter whether he votes or not, to the reaction of going and recording his vote. If we achieve the desired modifications, we have been successful in our objectives.

Objectives not Determined by Psychology. — What the objectives in education are at any given moment is not the outcome of psychology. Objectives are determined by the social philosophy of the age, by the actual needs and wishes of a given community or group. Whatever the objectives may be, whether desirable or undesirable, specific or vague, psychology can, however, give very decided help to the educator in attaining his aims. This is the special function of educational psychology.

Definition of Educational Psychology. — General Psychology is often defined as a study of the behavior of the individual in response to any and every kind of situation that life presents. If now, we concentrate on those situations in life that may be called specifically educational, many problems arise that are not ordinarily dealt with in general psychology. We may, therefore, define educational psychology as a study of the behavior of the individual in response to educational situations.

The Two Chief Divisions of Educational Psychology. — The starting point for all and any kind of education is the individual. It is the individual that we want to change. Hence the first big division of our subject is an intensive study of the individual. If we are to be good craftsmen, we must know the material with which we have to work. A good carpenter will know something about wood so that he may know how to use the different kinds for the various things he constructs. Some knowledge of the raw material with which he works is necessary to the successful worker. If he is lacking in such knowledge he will try to achieve impossibilities or he will bring forth monstrosities. We cannot make silken purses out of sows' ears, although sometimes we try to do so.

The second division of our study is the problem of modifying this raw material. What are the laws governing the modification of the individual? We know that we cannot make any changes that we wish. We cannot hope to make any kind of modification that our desires dictate or our fancy suggests. We may desire to make all people intelligent and educate them all up to an I.Q. of 150, but we know this to be impossible. Imagination can conceive of individuals equipped with radio receiving senses and radio transmitting mechanisms, but common sense knows that education can never effect such modifications. What are the changes that can be made? How may we best accomplish these changes? Perhaps some are easier to make at one time of life than others. What are the best stimuli or instruments to effect the changes we desire?

Original Nature and Its Modifications. — So we may divide the whole of Educational Psychology into two parts:

I. The Original Nature of Man.
II. The Modification of Original Nature.

And you will notice that this book is so divided into two parts. Educational Psychology tries to answer these two questions: 1. What is the nature of the material with which the educator has to deal? 2. How can this material be modified? As in any other science, these two divisions are arbitrary. They are made merely for the sake of convenience. The individual is never stationary or static. He is changing more or less all the time. Any and every stimulus to which he reacts is causing more modification of his original nature. The forces of environment are from birth continually working upon him and changing him. But note the fact that they are working upon *him* — upon a given type of material and they can only make such modifications as this material, of which he is made up, will allow. So we may very profitably study separately these two aspects, one at a time.

Original Nature. — If we watch the amazing mixture of reactions which individuals in general make to the countless objects in their environment, it would seem impossible to bring any order out of the chaos. But as we watch we note a certain similarity in reaction to certain phases of our environment. An object such as a hand or ball approaching the eyes suddenly generally causes the eyelids to wink. The sudden appearance of a large unknown object in the dark generally causes the individual to start and perhaps to tremble or even run, or make such reactions as would be ordinarily called fear reactions.

The presentation to an individual of a board with differently shaped cut-outs and a verbal or gestural stimulus of "go on," will generally lead to an attempt to fit the blocks in the board. And so we could continue with many other examples. This general similarity of response among individuals is in the first place due to the common inheritance of a certain type of physical body possessing a certain type of nervous system. It is because our nervous systems are all more or less alike that we find this general similarity of response. Of course, it might be argued that any of the sample reactions cited above might be modified, if the environment be changed. Prolonged practice may teach us to inhibit the winking reflex at the sudden approach of an object to the eyes. We might learn to run to a large unknown object seen approaching in the dark. We might learn to throw the board and blocks away whenever they are presented to us. But these are not the ordinary reactions that take place, and it is the ordinary reactions which arise in the ordinary environment of man, without specific attempt at modification, that we are going to consider his original endowment. For educational purposes it is this stock of reaction tendencies, which arise without specific training, that forms the foundation for further modification.

This native equipment of man can be conveniently divided up into four divisions:

1. Reflexes.
2. Instincts.
3. Emotions.
4. Capacities.

The first division containing the reflexes has to do with those reactions limited to various parts of the body which follow very regularly upon the presentation of a given stimulus. The winking reflex is a good example. The patellar or knee-jerk reflex is another. There are many reflexes. In education we are not concerned directly with them. The instinctive reactions are the more complex reactions made by the whole body. There are a great many of them and they are difficult to classify. Their significance for education is, however, great, because they form the basis for much of our conduct. Our character and to some extent our interests in life are largely determined by the strength or weakness of various groups of instinctive tendencies, and by the successful modification of these tendencies in various directions.

An emotion is an hereditary type of reaction involving changes of the bodily mechanism as a whole, but particularly of the visceral and glandular systems. The shock of the stimulus throws the organism into a more or less chaotic state for a brief period. This occasions the vague indefinite stirred-up feeling characteristic of the emotions. Emotions, as well as instinctive tendencies, are important for character development. The proper modification of our emotional tendencies leads to emotional stability amid the shocks of life. Poor or faulty modification of these tendencies results in emotional instability and its resultant character defects.

In the last division, namely the capacities, we are concerned mainly with reactions of an intellectual type. Individuals differ by original nature as to their ability to learn. With some individuals learning takes place quickly and easily and with others learning is slow and hard, in spite of similar opportunities and incentives. Perhaps there are special capacities for learning some things rather than others, such as words or numbers or music or mechanical things. More likely such differences are due to different interests based upon instinctive tendencies, or different amounts of sensitive acuity as in music, and undoubtedly many of these differences exhibited by children and adults are due to environmental factors influencing the child from early life. At any rate differences in the general ability to learn are obvious and such differences are all-important in education.

In this study of the original nature of the individual we shall note that the great advance made by modern psychology has been the attempt to describe in quantitative terms. If we can tell the educator how much general learning ability each child possesses he can plan more definitely for each child's education. With this amount of intelligence we can expect to make this amount of modification and so on. Hence an important part of our study will be to see what instruments of measurement have been devised and what they can accomplish. Not only in this field of general capacity, but also in the vaguer field of character qualities, we shall note the growth of quantitative work, and the help that this gives us in learning more about the original nature of the child.

The Modification of Original Nature. — Having studied the general nature of the individual, we must inquire how modifications of this original nature are brought about. Man is the most modifiable of all animals and modifications of his reaction tendencies are numerous. The school sets out very definitely to make certain types of modifications. So we are next confronted with the Laws of Learning, which attempt to describe how modifications in general are achieved. A study of these laws or principles will give us some insight into the general characteristics of learning and may give us hints as to how people may study most effectively. It may help the teacher to direct more economically the learning of his pupils.

The best way to understand how modifications of reactions take place is to have individuals learn different things in different ways and to record carefully their learning from day to day. When we do this we obtain practice curves and a study of such curves will tell about the rate at which modification or learning takes place for various things learned. We shall also find out something about the amount that can be learned and this will lead us on naturally to discuss what limits are set by the human organism as to the changes that can be made.

When changes or modifications have been made, when things have been learned, do they remain fixed? How permanent are they? This raises the question of forgetting and we shall have to study how forgetting takes place, how modifications change or tend to disappear when practice ceases.

We do not learn one thing at a time. We do not wait until we have learned one thing before beginning to learn another. We are always learning a great many things at the same time. In school a pupil is studying several subjects at the same time. Does the study of one subject have any influence upon another? What influence does one modification that is taking place have upon another modification that we are trying to effect? If we have already learned something, what influence does this have upon other things that we are now learning? We might call this reciprocal modification, the influence of one modification upon another. It is, of course, the important problem of transfer of training.

If we go on learning, practicing, trying to make modifications continuously without any pause for rest, we eventually become tired. Continuous work leads to fatigue. How long can we work efficiently before fatigue becomes noticeable? What effect does fatigue, mental or physical, have upon the rate and efficiency of our work. Closely allied to this problem is the influence of the loss of sleep upon work, and also the influence of drugs upon mental efficiency.

Throughout all this discussion of modifications of reactions or learning, we shall note how psychologists have devised various experiments to measure these modifications. But, finally, we shall turn to a study of the instruments which psychologists have devised to measure very specifically some of the chief modifications that the school is more particularly trying to effect. The school deliberately tries to build up a set of reactions towards printed words, which we call reading. It builds up another set of reactions towards numbers, which we call arithmetic, and so on for the other school subjects. How effective is the school in making these modifications? To measure these specific improvements made by the school there have arisen a number of measuring instruments called educational tests and they are of extreme importance for the educator. They are objective measures of the modifications of original nature, which the school is attempting to make.

Summary

This chapter has given a brief outline of the field of educational psychology. If you glance over the chapter headings in the Table of Contents, you will see that the chapters of the book follow this outline of the contents of the book. Let us put in schematic form the field of our subject, so that we may keep it definitely before us throughout our study.

I. Original Nature.

1. *Reflexes.* Of slight educational significance.
2. *Instincts.* More complicated reactions. Important for interests and character.
3. *Emotions.* Important for character development.
4. *Capacities.* General term for more specifically intellectual reactions. Measurement of these by means of intelligence tests.

II. Modification of Original Nature.

1. *Laws of Modification.* The laws of learning. The way changes are made.
2. *Making Modifications.* Practice. Amount, rate and limits. Economical ways of learning.

3. *Permanence of Modifications.* Forgetting.
4. *Reciprocal Modification.* Transfer of Training.
5. *Continuous Modification.* Fatigue.
6. *Measurement of Modifications in Education.* — Tests of school subjects.

Review

True-False Statements

As a review of what you have read, mark the following statements true or false. If you have any doubt, go back to the text.

1. We inherit a certain type of nervous system which determines the kinds of responses we can make.
2. The modifiability of original tendencies is a measure of the educability of the individual.
3. The primary purpose of educational psychology is to set up certain aims and ideals towards which educators should strive.
4. Educational practice is based upon the fact that most original tendencies are modifiable.
5. All methods of education should build on the original tendencies of the child.
6. Education in general attempts to preserve intact the original and natural tendencies of the child.
7. Emotions and instincts are the basic elements upon which character is built.
8. There is no limit to the possible modifications of human behavior.
9. A man is the result of his original nature and his environment.
10. Educational tests are quantitative measures of the modifications of original nature achieved by the school.

PART I—ORIGINAL NATURE

CHAPTER II

THE ORIGINAL TENDENCIES

Inheritance of Definite Nervous System. — Man inherits from his parents a certain type of body which includes a certain type of nervous system. This means that he inherits a given set of receiving organs for receiving impressions from the external world, a central nervous system, and also a given set of end organs, muscles, glands, and the like, by means of which he may respond to the impressions which he receives from the outside. Because of this structure which man inherits, he tends to respond to certain stimuli in a given way. Animals respond to stimuli in general somewhat differently from man. The greater the difference between the nervous system of an animal and that of man, the greater is the difference between the responses of the man and the animal likely to be.

Different Kinds of Original Tendencies. — These general tendencies to respond in a given way to a given stimulus are called original tendencies. As we saw in Chapter I, some of these tendencies are called reflexes, when the response is limited to a very small part of the body. The larger responses, resulting in a more complex type of activity, issuing in a rather vague set of movements are generally called instincts. Still vaguer tendencies to respond are sometimes called capacities. These do not result in obvious physical movements. They may be provisionally described as tendencies to think, to remember, to learn and the like. Responses which result in an internal general "stirring-up" of the organism and do not lead to any very definite set of movements are called emotions.

No Sharp Division between these Tendencies. — There is no sharp dividing line between these various tendencies to reaction. We must be careful not to let the names for these various tendencies become magnified into powers or explanatory causes of our actions. If we say that speech is an instinct, we are in danger of assuming that this is an adequate explanation of how an individual learns to talk and to rest content with such an explanation. If we say that curiosity is an instinct, it affords an easy explanation of all the new things we like to try and do, and we tend to ignore the fact that there are many new things we do not like to try or do. For this reason some psychologists think it better to abandon the name of instincts altogether. What we inherit is merely a vague unrest, a vague and indefinite tendency to respond. The modes of unrest depend upon the presence of specific organs and mechanisms. The particular responses that result in any case depend upon the environment that surrounds us. And from the very beginning this environment gradually conditions or influences the specific responses that are made. From the very beginning, therefore, there is an active interplay between the organism and the environment and no original tendency or instinct stands out clearly and independently of the environment.

Continuum of Original and Acquired Reactions. — Original and acquired reactions are merely two names given to the two extremes of a continuous series of reactions. There can be no dividing line between them. If we were to list a long series of responses to situations, beginning with those that are most predictable and going gradually over to those less and less predictable, we would have those reactions ordinarily called reflexes and instincts at the one end and those called acquired reactions at the other. Let us try to make such a list, putting down the stimuli at one side and the probable reaction or reactions on the other.

Stimulus	*Response*
increase of illumination	contraction of pupil of eye
sudden loss of support	clutching, reaching, fear
tap on patellar tendon	knee jerk
movements of stomach walls	hunger
bitter substance in mouth	spitting
sight of novel object	reaching, handling
calling of one's name	answering in some way
word association to "king"	"queen"
hear telephone ring	pick up receiver
two plus two	four
see acquaintance on street	smile, nod, etc.
receive invitation	accept, reject
see postage stamps	collect, ignore, etc.
receive some money	spend, save, etc.
look at a picture	various thoughts

Original Tendencies very Vague. — As we read down the column of stimuli, it is obvious that we are less able to predict the reaction that would take place in human beings in general. Even from the very start our probable responses assume a healthy individual, because the pupillary reflex and the patellar reflex are impaired in certain diseases, and some idiots do not spit out bitter substances placed in the mouth. Only with a certain percentage of educated individuals will the response "queen" be given to the stimulus "king" (see page 56). Not all of us collect postage stamps and what we may think, say or feel when looking at a given picture is very difficult to predict.

All that we mean by an original tendency to reaction, then, must be that a given stimulus is more likely to call forth one kind of a response rather than another, because of the particular make-up of our organism. These original tendencies will be vague and general and subject to much modification.

Lists of Instinctive Tendencies. — Because the human being possesses a highly complicated nervous system, there are a great number of these instinctive tendencies to action. Almost any text-book of psychology discusses them thoroughly and many psychologists have attempted to make lists of instincts or groups of instinctive tendencies. There is naturally great difference of opinion as to just which reactions are instinctive and as to how those that are considered instinctive should be grouped. We shall content ourselves here with giving two lists of instincts, one suggested by Thorndike and the other by Woodworth.

ORIGINAL TENDENCIES IN HUMAN BEHAVIOR

Thorndike

1. Sensitivities
2. Attentiveness
3. Bodily control
4. Food, habitation
5. Fear, fighting
6. Social instincts, gregariousness
7. Manipulation
   Visualization
   Vocalization
8. Expression of emotions
9. Learning, remembering

Woodworth

1. Responses to organic needs
   Thirst, hunger, breathing, heat and cold, shrinking from injury, crying, fatigue, sleep

2. Responses to other persons
   Herd or gregariousness
   Mating or sex
   Parental or mothering

3. Non-specific or play instincts
   Kicking, running, etc.
   Locomotion
   Vocalization
   Manipulation, exploration
   Laughter
   Fighting
   Self-assertion
   Submission, docility

It is not necessary for the student of education to memorize these lists or any other lists proposed by other writers. Naturally psychologists will differ greatly in their descriptions of these vague general tendencies to respond.

The Emotional Responses. — In addition to these modes of response, which are generally called instinctive,

FIGURE 1. — What emotions or feelings are indicated by these pictures? (From Feleky's *Feelings and Emotions*.)

there are also other responses of a general nature, which we call emotional responses. These responses are characterized by significant physiological changes in the organism. There is increased activity in some glands and restriction of activity in others. Thus in anger the secretion of the gastric juice into the stomach is restricted, while the activity of the adrenal glands is increased. And so it is with other emotional states. They are responses of the organism to stimuli in its environment, and at the same time they are physiological urges to further action. When something makes us angry, we may go on to fight or to overcome it; when something makes us afraid, we experience an emotion of fear and this is preparatory to further reactions of flight or hiding or crying or shouting for help or braving it out and so on according to the environment and our previous training.

Like the so-called instinctive modes of response, these emotional modes of response are very general and vague. A burst of tears may be the result of intense joy or anger or fear. In ordinary life we can interpret the emotional expressions of others fairly well, if we know the circumstances leading up to the expression. In watching a moving picture, we also have little difficulty because we know the story, and the actors in addition generally exaggerate certain conventionalized modes of expression so as to leave us in little doubt about the emotions they are supposed to be experiencing.

But if we do not know what led up to an emotional outburst, and if we are not familiar with the individual's usual modes of expression it is by no means so easy. Study the pictures in Figure 1, and jot down underneath what emotion you think is portrayed. When you have done this turn to the end of the chapter and check your results. Emotional modes of response, just like any other modes of response, are modified by our social environment. Upon the original vague modes of response are built different types of expression to different types of situations, and there are considerable individual differences in these modes of expression.

Original Tendencies Determine General Form of Society. — The original tendencies determine the general background of our civilization and, therefore, of our educational aims. The fact that man is a social animal, that he is gregarious and tends to feel satisfaction when living and working with others of his species, is the basis of our present form of civilization with its cities and villages, its family groups and societies. In this he resembles gregarious animals such as the buffalo or the sheep, rather than solitary animals like the tiger or the leopard. If man had been predominantly a solitary animal, the form of our civilization and the type of our educational system would undoubtedly have had a different sort of evolution than the one we are familiar with at present.

Manipulation. — Manipulation is mentioned by most writers as one of the important original tendencies. By manipulation is meant the general tendency to touch with the hands or feet objects in the immediate environment of the individual. We reach forward to feel and handle any strange or new object, unless we have been taught not to do so by previous experience with certain classes of objects. This general tendency to manipulate or handle is one of great educational significance. By means of it the baby learns about countless things in its environment. Some of these things cause pain or annoyance and it learns to avoid such things in the future. All through life this tendency persists to a greater or less degree. We feel we do not "know" an object unless we have handled and manipulated it. The desirability for doing things in school rather than merely reading or hearing about them is obvious. Merely reading about the Eskimo in the second grade may teach the children something valuable, but if they are encouraged to make an Eskimo village on their sand table, with igloos and boats and all the rest, they will be vastly more interested and they will undoubtedly "know" more about Eskimos when they have done this. So in the higher grades and in the university, the laboratory and the workshop should supplement the recitation, the lecture and the book, in order that we may make use of and foster this basic instinct of manipulation. Perhaps the individuals who have inherited a very strong tendency to manipulation and who have been encouraged to develop this tendency, are those who have become our great inventors and our master mechanics.

If this tendency is fostered, encouraged and rewarded, it will expand and flourish in many ways. A social or educational system may do this. The history of the United States shows a society in which the manipulative tendency and habits built thereon have been highly prized and highly rewarded. This was not only the outcome of the pioneer life, in general, but also the result of the type of democracy which exalted the farmer and the mechanic somewhat at the expense of the mental worker. The great industrial and mechanical transformation in the nineteenth century gave further encouragement to the manipulative tendencies of man. The mechanical genius and inventor is to-day the idol of our youth. On the other hand there are civilizations and there have been periods in history where tendencies to manipulation and habits built thereon have been repressed and utterly despised. To work with the hands was not fitting for a gentleman. The ideal in life may be contemplation and not action.

Vocalization. — The original tendency to vocalization is at the basis of all speech. This tendency is seen very early in the babbling and cooing of the infant. There is evident satisfaction resulting from the sounds made.

Certain sounds will be repeated over and over again. The child soon learns that the making of certain kinds of sounds brings very definite satisfactions, such as more attention from the mother, more food, and the like, and so these sounds tend to be repeated and retained and other sounds tend to be neglected and forgotten. The child does not learn by imitation. We have to teach the child to imitate, if we want him to imitate. Indeed in the early stages of language development, it might more truly be said that the fond parent imitates the gurgling and cooing of the child. So much so, that in some cases children concoct words of their own which parents perpetuate, and in many cases the baby talk of the infant is carried over into later childhood much to the consternation of the parent. Indeed we might readily build up a new language out of the babbling of the infant by rewarding some sounds and repressing others, by attaching certain sounds to certain situations in the same way as we learn everything else in life.

Vocalization begins very early. The cry of the new born infant is one of the first reactions after birth. "The respiratory tricks, the cooing and babbling of even the young infant are a preparation for language," according to Gesell (1). This author reports a twenty-four hour count of the vocal reactions of a six-months-old child. Figure 2 shows the number of times each separate sound was used. Da, for example, was spoken 63 times, ngrr occurred 21 times, and so on. Altogether 104 separate vocal reactions were made during the twenty-four-hour period. Considering how much time a six-months' baby usually spends in sleep, this number of vocal reactions emphasizes the importance of this tendency.

Vocalization is the basis for all language. What specific language will be learned by a child depends, of course, upon its environment. To the young infant be-

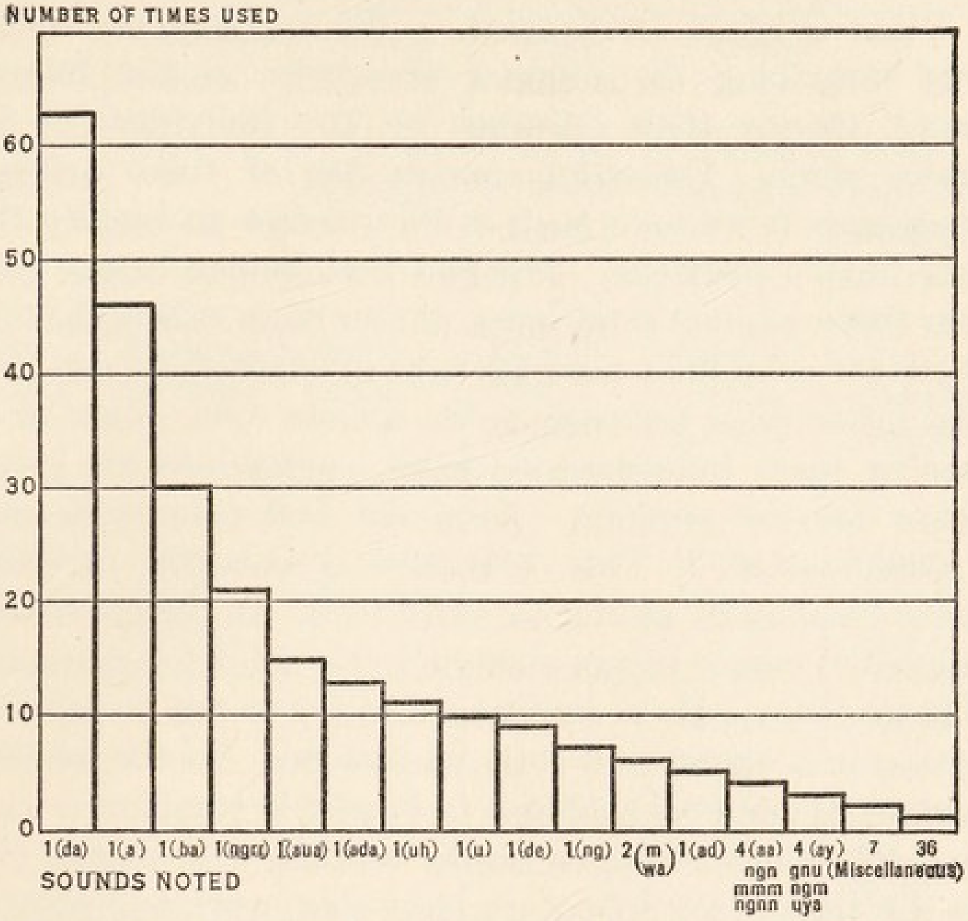


FIGURE 2. — Vocalization Chart of L. M., Age Six Months. Sixty-four different sounds were distinguished and their frequency is shown in the diagram. (From Gesell's *The Mental Growth of the Pre-School Child.*)

ginning to vocalize no language possesses any terrors. It would just as readily learn Chinese, or Polish or Greek as English. Only after we have learned one set of language habits, does the acquisition of a new set present any difficulties. When the child enters school this tendency to vocalize is used in the story-telling and singing of the kindergarten, in the discussions about things and in dramatics and so on in the upper grades. Attempts to repress this tendency in school by rules against talking, by attempting to enforce silence are generally futile. The wise teacher will use this very strong tendency as a powerful ally rather than struggle incessantly against it as an enemy.

Other Original Tendencies. — We might go on in this way describing the original tendencies of the human being, tracing their influence on the individual as he grows older. Education makes use of these original tendencies in various ways in its attempt to modify the individual's reactions. The fact that human beings possess these original tendencies, rather than others, has on the other hand influenced the type of educational methods and institutions common to the human race. Instead of tracing these influences in detail, suggestions are given below for the student. Take the first column headed "manipulation." This tendency is enlarged in early home training by giving the child toys. In the pre-school education many things surround the child to encourage this tendency. He is encouraged to try to put on his own coat, use a spoon and fork and so on. In the kindergarten, educational material to handle is common — cutting out things, stringing beads, building with blocks, etc. In the elementary school we have shop work and sewing, in the high school and college we have our various shops and laboratories. The diagram on the following page is merely suggestive, the student should work out others for himself.

Original Tendencies Influence our Conduct. — All through our lives these original tendencies, greatly modified in many cases, are at work. Advertisers make skillful appeals to them. We all wish for the approval of our fellows more or less. We shrink from ridicule. We want to be as good as the other fellow, and so we are greatly moved by the harrowing story of the young woman whose husband lost forever his chance to become "general manager" because he did not know the proper way to eat olives at the great banquet, and so we rush wildly to our booksellers for a copy of The Great Book on Etiquette lest we should suffer a like fate. We denounce kings and potentates and all their evil ways, but thrill

	Manipulation	*Bodily Control*	*Gregariousness*	*Vocalization*	*Fear*
Early home	Toys	Crawling Walking	Playing with parents and sibs	Crowing Gurgling	Disapproval of mother
Pre-school	Toys, useful things	Stairs Slides	Playing with others around	Talking to other children	Disapproval of teacher
Kindergarten	Educational material	Marching Dancing	Plays	Stories Songs	Disapproval of teacher
Elementary	Shop Sewing	Games Stunts	Games	Poems Songs Dramatics	Disapproval of teacher and group
High school	Shop Laboratory	Athletics Dancing	Organized Games Societies	Dramatics Debates	Group disapproval
College	Laboratory	Athletics	Organizations of all sorts	Dramatics Debates Speeches	Group disapproval

with a glow of pride when we are elevated to the high position of Grand Master of the High and Sacred Order of Elephants or Monkeys. We scorn the titles and decorations of an effete European aristocracy, but we eagerly decorate our bosoms with and suspend from various parts of our anatomy pins and keys and mystic symbols to advertise our greatness and mark us off from the common herd.

These general examples are given so that the student may realize how original tendencies form the background of our lives. We must be careful not to imagine that one instinct works alone and independently of all the rest. When we suggested that a particularly strong tendency to manipulation may be one of the factors producing the great inventor, we must remember the presence and influence of countless other tendencies. The inheritance of a great capacity to learn would also be necessary, and a suitable environment that would encourage and stimulate all these inherited tendencies is of course pre-supposed. The lack of a suitable environment, the emphasis of a non-active type of education might easily modify or weaken the instinctive tendency towards manipulation, and a great inventor would not be produced.

The Modifiability of our Original Equipment. — It is important to remember how easily modifiable these original tendencies are, some undoubtedly more so than others. Man, as contrasted with other animals, may be thought of as possessing the most loosely organized, the most readily modifiable, set of original tendencies. This makes man the great learner among animals. This fact of plasticity or modifiability makes man more susceptible to his environment than other animals. This modifiability is due to the plasticity of his nervous system. The connections are loosely organized, and, therefore, there is great possibility of reorganization and modification. This plasticity is, of course, greatest at birth and gradually decreases with growth. The adult is not so plastic or modifiable as the child. Note also the long period of childhood in man as contrasted with the relatively short period of childhood in most animals. The shorter the period of childhood, the less opportunity is there for modification. In an insect we have a very highly organized set of instinctive tendencies with a very short or complete absence of a period of childhood. Certain stimuli are responded to from the very beginning in a definite way and almost no modification is possible throughout life. The insect is mature from the very beginning, busy at all times responding definitely to the appropriate stimuli. Man grows slowly; he is not busy or intent upon a task every moment of the day. He is, as Woodworth so well expresses it, "the most pottering, the most hem-and-hawing of all animals." It is because he potters that he discovers various ways of doing things. It is because he "hems-and-haws" that he discovers an infinite variety of modes of expression, and develops complicated languages in place of the restricted means of vocal expression of the dog or monkey or bird.

The Vagueness and Indefiniteness of Our Original Equipment. — So modifiable are our hereditary modes of response and so early in life are they susceptible to modification, that some psychologists question whether there are really any instincts at all. They believe that the organism exhibits at birth a great number of bodily movements which are unorganized and chaotic. These are the units of reaction, and out of these, by means of the stimuli of the environment playing upon the organism, various reaction systems are built up. Whether this be the case or whether we inherit more specific modes of response as we have previously indicated, the important fact for the student is to remember the great possibility of modification that exists from the very beginning of the life of the organism. Upon this fact of modifiability all our hope of great educational improvement depends.

Few Emotional Responses. — In a similar way, the number of original emotional modes of response have been limited to three, namely fear, rage and love.

In work with very young children, after trying many different stimuli, three types of responses seem to be common. These have been called by Watson fear, love and rage.

Fear: Stimuli		1. Sudden removal of support from under child
		2. Loud sounds
		3. Sudden disturbance when falling asleep
	Responses	1. Sudden catching of breath
		2. Random clutching with hands
		3. Puckering of the lips
		4. Crying
Rage: Stimuli		1. Hampering movements
		2. Holding arms by side
		3. Holding head
	Responses	1. Crying and screaming
		2. Slashing and striking movements
Love: Stimuli		1. Stroking
		2. Tickling
		3. Rocking
		4. Patting
	Responses	1. Smiling
		2. Gurgling
		3. Cooing
		4. Extension of arms (embrace)

These may be the chief situations calling forth emotional reactions in the very young child. Undoubtedly as the child grows older other situations will call forth original emotional tendencies to respond. But from the very first months of life other stimuli become attached to the original modes of response, and the modes of response themselves become modified. Many of the complicated emotional reactions of the adult are to be thought of as modifications and mixtures of these three simple types.

How do these modifications arise? If there are only a few specific situations which call forth the emotion of fear in early life, how do we account for the numerous fear reactions which the child shows as he grows older? A study of some experimental work done with babies by Watson (4) may make this clearer, and at the same time have important bearings upon our understanding of the development of character qualities in general in the individual.

Conditioned Emotional Reactions. — At nine months of age the child started violently at the sudden sound of striking a steel bar behind his back. The second stimulation produced a violent start again and in addition the lips began to pucker and tremble. A third stimulation caused the child to cry and show the common expressions of fear. We may consider these reactions as native or original with respect to the stimulus of a sudden loud unrecognized sound. The question now is, are these fear reactions also called forth by many other stimuli when presented to the child for the first time. Evidently not, for the infant was confronted suddenly and for the first time with a white rat, a rabbit, a dog, a monkey, masks with and without hair, cotton wool, burning newspapers, etc. To most of these stimuli the instinctive response was one of reaching out to touch and manipulate. None of them called forth the crying, shrinking responses of fear, which were elicited by the loud sound. How, then, does the emotional reaction of fear become attached to other objects? How can we shift the instinctive reaction called forth by one stimulus over to another stimulus which does not by original nature elicit it in the first place?

Let us continue our description of the experiment. The white rat was again presented to the infant. As before he began to reach for the rat and just as his hand touched the rat, the steel bar was struck immediately behind his head. He jumped violently and fell forward. This was done twice. A few days later the rat was presented again and he started to reach forward, when the rat touched his hand, the hand was immediately withdrawn. Five more presentations of the rat with the loud sound were given, and there occurred an increasing tendency for the child to show fear and to cry. After this the rat was presented alone, without any sound stimulus. Immediately the infant began to cry and to crawl away as rapidly as possible.

We see, therefore, how a given response (fear) attached by original nature to a given stimulus (loud sound) has been shifted over to another stimulus (rat). We may also say that a given response (manipulation), attached by original nature to a given stimulus (rat), has now been modified or changed into a different response (crying and crawling away).

Transfer of Fear to Other Stimuli. — The experiment we are describing continues to show how the fear response called forth by the rat was also called forth by a rabbit when presented to the child. We must remember that the very first response to the rabbit, as to the rat, was a manipulation response. Now after the reaction to the rat has been changed to a fear reaction, the rabbit also calls forth a fear reaction. It is not necessary to go through a series of experiments with simultaneous sound and rabbit to change the original manipulation reaction. There is evidently sufficient similarity between the rat and the rabbit to effect a transfer from the one to the other. There are evidently enough identities between the two situations to cause a response learned in connection with the first situation to transfer over to a second similar situation. This is an important factor in learning and we will take it up later in detail in its wider educational significance. Here we may simply note that once the child has been taught to fear a rat, he has also been taught to fear similar animals, such as a rabbit. The experiment goes on to to show that this was true also, to some extent, for a dog and a fur seal coat, but not for a package of cotton wool, a Santa Claus mask or the hair on the head of the experimenters. Furthermore it was shown that after an interval of thirty-one days, the stimulus of the rat still called forth the fear reaction.

The Process of Unconditioning. — Now let us raise the question as to whether we can reverse the process described in the above experiment. Suppose a child has already learned to fear a rabbit, can we employ the same procedure in order to change the fear response to the rabbit? In this case we must reverse the process and associate with the fear object a definite stimulus which arouses a pleasant reaction. What kind of a stimulus is potent in calling forth a pleasant reaction? Obviously the sight of food when hungry is by original nature attached closely to reaching and touching reactions and later, when the food is being eaten, there are strong feelings of satisfaction and pleasure.

Another experimenter, Jones (2), has shown that in this way a child who exhibited fear of a rabbit could be taught to like the rabbit. In this case the child was 2 years and 10 months old and at the beginning of the experiment had already formed fear reactions (conditioned reactions) to a rabbit, white rat, fur coat, etc. The "direct conditioning" consisted in introducing the rabbit to the child while he was being given food which he liked. Through the presence of the pleasant stimulus (food) whenever the rabbit was shown, the fear was eliminated gradually in favor of a positive response. Steps in the degree of toleration listed by the experimenter show very nicely how the fear was overcome:

1. Rabbit anywhere in room in cage causes fear reactions.
2. Rabbit 12 feet away in cage tolerated.
3. Rabbit 4 feet away in cage tolerated.
4. Rabbit 3 feet away in cage tolerated.
5. Rabbit close in cage tolerated.
6. Rabbit free in room tolerated.
7. Rabbit touched when experimenter holds it.
8. Rabbit touched when free in room.
9. Rabbit defied by spitting at it, throwing things at it.
10. Rabbit allowed on tray of high chair.
11. Squats in defenseless position beside rabbit.

12. Helps experimenter to carry rabbit.
13. Holds rabbit on lap.
14. Stays alone in room with rabbit.
15. Allows rabbit to play in pen with him.
16. Fondles rabbit affectionately.
17. Lets rabbit nibble his fingers.

This list showing the changes in reaction tells in brief how the baby became more and more accustomed to the rabbit. The experimenter did what many a wise mother might well do in order to get her child over a fear for an animal. At no time was the child suddenly shocked by the near presence of the animal. The procedure sounds like good common sense. It is well, however, for us to have the procedure worked out systematically, so that we may understand fully how gradually such a resistance to a fear is built up. Other mothers or teachers not so wise might have suddenly thrust the rabbit upon the child, compelled him to hold it and touch it, and by so doing increased rather than diminished the fear. Compulsion, with reference to things a child dislikes, probably very often increases the dislike rather than lessens it. Association of the disagreeable with agreeable associates or consequences may be the most effective means of removing the dislike or fear of a given situation.

Diagram of Conditioning. — The changes in reaction which these experiments have demonstrated are technically known as conditioning and unconditioning. The conditioned response is the changed or modified response. What has been happening in these experiments may be represented diagrammatically, where S stands for stimulus and R for response and → for "is followed by."

Conditioning. First Experiment.

S. 1. Loud Noise	→	R. 1. Crying (original tendency)
S. 2. Rat	→	R. 2. Manipulation (original tendency)
S. 1. + S. 2.	→	R. 1. Repeated seven times
S. 2.	→	R. 1. Original tendency conditioned

Transfer.

S. 3. Rabbit	→	R. 1.
S. 4. Fur Coat	→	R. 1.

No Transfer.

S. 5. Cotton Wool	→	R. 2.
S. 6. Mask	→	R. 2.

Permanence.

S. 2.	→	R. 1. (After 31 days.) How long will it last?

Unconditioning. Second Experiment.

S. 2. Rabbit	→	R. 1. Fear (conditioned response)
S. 7. Playthings	→	R. 2. Manipulation (original tendency)
S. 8. Food	→	R. 2. Handling, satisfaction (original tendency)
S. 2. + S. 7.	→	R. 2. Several times
S. 2. + S. 8.	→	R. 2. Several times
S. 2. Rabbit	→	R. 2. Manipulation

We have described these experiments in modifying original tendencies in some detail because they illustrate admirably how very modifiable our original tendencies are, and at the same time they indicate how this modification begins at a very early age.

Other Conditioning Experiments. — These experiments in conditioning or changing a response from one situation to another have been stimulated by the Russian physiologist, Pavlov, and his pupils. The classical experiment is the measurement of the flow of saliva in the dog as illustrated in Figure 3. The sight or smell of food increases the flow of saliva. If food is always presented along with another stimulus, the sound of a bell or a colored light, then in course of time the sound of the bell or the colored light will act as an adequate stimulus for the salivary flow.

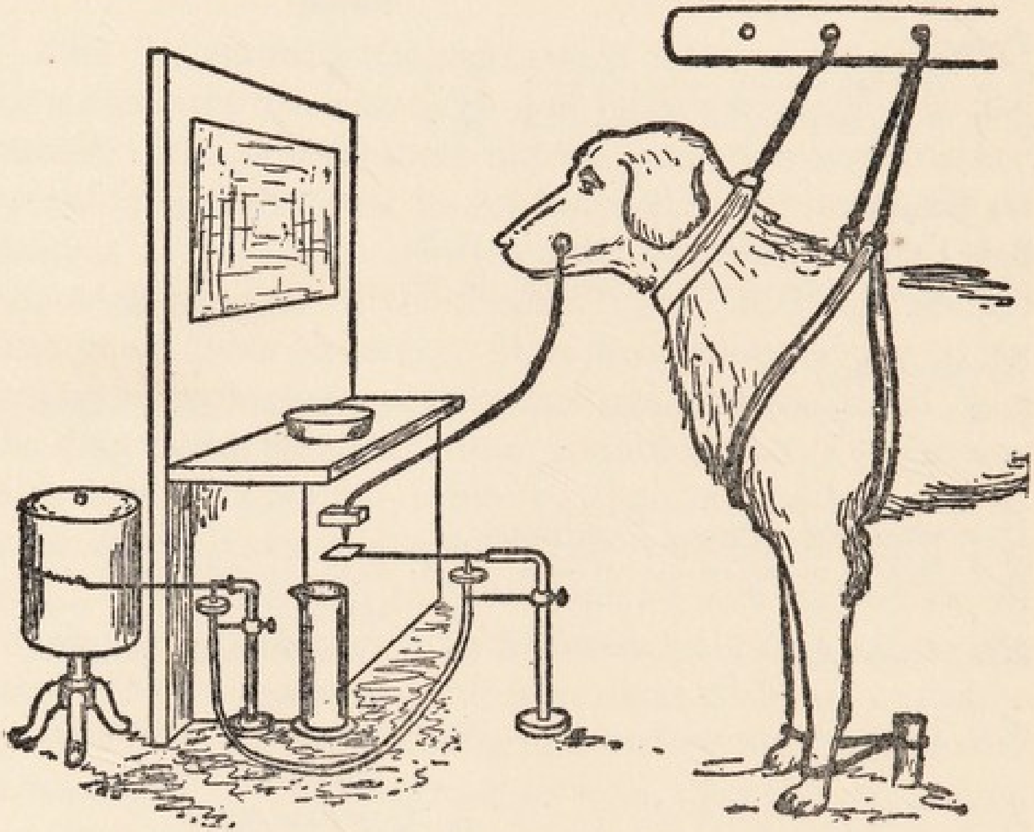


FIGURE 3. — Pavlov's Conditioned Reflex in the Dog.

Many other conditioning experiments have been carried out with animals, children and adults. Such experimentation suggests the ways in which children become conditioned in everyday life. A slamming door as the child gropes its way up a dark stair may be the beginning of fear of the dark. A four-year-old child saw a snake and ran in delight to her father to bring him to see the pretty animal. As they went back to look for the snake hand in hand, coming upon it suddenly, the swish of the snake in the long grass made them both start back, the father reënforcing the fear of the child, and so may have been started a lifelong fear of snakes. Up to her fifth year a little girl had never objected to eating chicken, but after returning from the country she declared she did not like the taste and from that time on always refused it. She had seen a chicken killed while on her country visit. Many years after she still "hated chicken," although she had long forgotten all about the killing of the chicken. As a matter of fact the taste of chicken was not distasteful to her, because she ate it once when she thought it was something else. The transfer effect of the chicken incident rapidly spread, so that she soon "hated" turkey, duck, goose and other birds.

Importance for Education. — The fact that our hates, loves, fears and the like are being formed at an early age may be of extreme importance to education in general and character education in particular. We are coming to believe that a great deal of basic training can be achieved in the pre-school age. At present society does not see fit to compel children to enter school before the age of six. Kindergarten education has extended rapidly and many 4 and 5 year olds are to be found in our schools. Perhaps more deliberate conscious education is desirable even in earlier life. It may make a great difference to a child's character and temperament, if he can be freed from a great number of the fears, which accumulate in the ordinary uncontrolled environment; if he can be made to love a great many of the things which originally call forth neutral or negative responses. Original tendencies to manipulation, vocalization and the like may be extended or restricted, they may be turned into desirable channels more easily at an early age than in later childhood.

The understanding of how emotional reactions become conditioned or associated with situations which originally did not call out an emotional response is not only of value for early childhood, but it may also help us in many school situations. No doubt many of us "hate" arithmetic and mathematics and, therefore, avoid them, and, therefore, are not very good at them, because somewhere in our school career disagreeable emotions were aroused by the teacher or by circumstances surrounding the lessons. The fear or dislike became attached to the arithmetic and in course of time a habit of reacting in this manner grew up. In studying some high school children who were finding algebra very difficult, one investigator (Symonds, 3) found that several of them reported a dislike for arithmetic dating back to their elementary school experience. "Always hated arithmetic"; "cannot grasp mathematics" are some of the expressions the students themselves used. In a similar way, we may be able to understand some hatreds or fears of spelling. The spelling lesson may have brought ridicule upon the child and an emotional response of fear or anger is aroused and attached to the subject itself. Similarly our great likes and loves for certain subjects may arise because of the agreeable reactions called forth by the circumstances surrounding the lesson. Perhaps we "love" geography and history, because we started in a sandpile.

Of course, these reasons for ability or inability to master a subject are not the only ones at work. The general capacities of the student are important, as are also the methods of learning, but these are factors which will be treated at length later on.

Summary

1. The inheritance of the same type of body and nervous system leads to common modes of response among individuals of the same species.
2. These original modes of response are called instincts, emotions, reflexes, capacities, etc.
3. There is no sharp division between these different original tendencies, nor can original reactions be sharply differentiated from acquired ones.

4. All reactions are more or less acquired, but all are based upon the original modes of response inherited by the individual.
5. Lists of original tendencies made by psychologists vary very much, because our original tendencies are very vague and very modifiable.
6. Emotional modes of response are also very general and unless very extreme are difficult to interpret without previous knowledge of what led up to the emotional response.
7. The original tendencies, nevertheless, are important as they form the basis of our general form of society.
8. Manipulation is seen very early in the child and can be made use of at all stages of the educational process.
9. Vocalization is the origin of all language. It begins very early and is made use of in many ways all through school.
10. Many other original tendencies can be made use of by the teacher and they should be consciously so used.
11. The original tendencies are more modifiable in man than in other animals and hence man is the great learner. The plasticity of his nervous system and the long period of childhood give great opportunity for education.
12. The process of conditioning is one way by means of which we learn many things. This begins at a very early age and many of our fears, likes and dislikes are undoubtedly formed by this means in early life.
13. Hence the value of deliberate education in early life. The importance of the early formation of right habits is the psychological justification of the nursery school and other forms of pre-school education.

Pictures of Emotional Expression

The nine photographs in Figure 1 are supposed to depict the following emotional states:

1. Intellectual attention
2. Agreeable surprise
3. Horror
4. Contempt
5. Sneering
6. Laughter
7. Terror
8. Rage
9. Hate

Review

True-False Statements

As a review mark the following statements true or false:

1. Manipulation can be called an instinctive tendency to react.
2. Three basic emotional tendencies common to all infants are fear, rage and love.
3. Man has a more definitely organized system of original responses than most animals.
4. All instinctive tendencies are present and active during the first week of life.
5. It is a fair assumption that all instinctive behavior is useful at the present time.
6. Conditioning original tendencies always works for the good of the individual.
7. The best original tendency to use for disciplinary purposes in school is fear.
8. The toys of the nursery school appeal to the original tendency of manipulation.
9. Reactions conditioned in early childhood cannot be unconditioned.
10. The long childhood of man allows much opportunity for his original tendencies to be modified.
11. Man is originally attentive to sudden changes and sharp contrasts in his environment.
12. An original response may be transferred to other stimuli which did not previously arouse such a response.
13. You must teach a baby to walk if you want it to walk.
14. Many likes and dislikes of school subjects are probably conditioned responses.
15. When two situations producing conflicting responses are repeated simultaneously several times, the stronger response tends to eliminate the weaker and become the established response for either or both situations.
16. Fear of snakes is probably an original tendency.
17. We find bitter things disagreeable because our original nature tells us that they are harmful.
18. Instinctively he reached for the telephone to report the fire.

19. Hampering or interrupting the individual in what he is set to do almost always causes anger or rage.
20. The first few years of a child's life are probably extremely important for the formation of basic likes and dislikes.
21. Underline the best method of removing fear:
    (1) verbal appeal; (2) disuse; (3) direct conditioning; (4) repression.
22. Complete this sentence:
    Capacities, like reflexes and instincts, are inherited, but they are more ——— and more ———.

Advised Readings

Woodworth, R. S.: *Psychology*, New York, 1921, Chapters 6, 7, and 8.

Thorndike, E. L.: *Educational Psychology: Briefer Course*, New York, Teachers College, 1914, Chapters 1, 2 and 3.

Watson, J. B.: *Psychology, from the Standpoint of a Behaviorist*, Philadelphia, 1919, Chapters 6 and 7.

References

1. Gesell, A.: *The Mental Growth of the Pre-School Child*, New York, 1925.
2. Jones, M. C.: "A Laboratory Study of Fear: The Case of Peter," *Pedagogical Seminary*, 1924, Volume 31, pages 308–315.
3. Symonds, P. M.: "Special Disability in Algebra," *Teachers College Contributions to Education*, No. 132, Columbia University, 1923.
4. Watson, J. B., and Rayner, R.: "Conditioned Emotional Reactions," *Journal of Experimental Psychology*, Volume 3, No. 1, February, 1920, pages 1–14.

CHAPTER III

THE FUNCTIONING OF ORIGINAL TENDENCIES

In this chapter we shall describe the functioning of the original tendencies in the life of the individual, and what happens when the original modes of response are not permitted to take their normal course. We include both the emotions and the instincts in our general expression "original modes of response."

Conflict. — Let us begin with the fact of conflict. The infant responds to its environment directly by means of such original modes of response as it possesses. When stimulated to anger it screams and strikes at the offending object, when a thing pleases it the tendency is to grab and pull towards and probably put in the mouth. No one would think of blaming the child for such conduct. Indeed to most adults such behavior seems rather amusing. Similar behavior in an older child or adult is, however, not amusing. We expect modifications of these tendencies to take place as the child grows older and if they do not, we censure them. Now, all through life, tendencies to respond directly as the infant does, in accordance with our original modes of response, exist within us. We are taught to suppress them, to alter them, to turn them into other channels. When aroused to anger we very frequently feel a tendency to strike out or scream, but we realize that such conduct would cause disastrous results. But the tendency is there, nevertheless, stirring us up, seeking some means of expression. Some outlet seems to be necessary and so we gradually learn to modify the original modes of response and to seek other modes of expression.

Extroversion. — When the outlet is found in a reaction towards something in our external environment, we call it extroversion, a turning to the outside. The teacher reprimands the child, justly or unjustly. The child feels that he is being treated unjustly, because the reason for the reprimand is not clear to him or beyond his comprehension. He feels enraged, ready to strike and scream. But previous experience has taught him that such a reaction will only cause more trouble and unpleasantness. Hence he represses the original responses and turns them into making unjust remarks to others about the teacher, refusing to do his work, bullying smaller boys, and in extreme and longstanding cases, perhaps, playing truant, organizing a gang for the purpose of stealing or committing vandalism of some kind or other. It is the individual's way of justifying himself to himself. It is one way of maintaining his own self-respect. If joy and pleasure and success do not come legitimately to them in the course of their normal school activities, some children will find these devious paths to obtain the satisfactions to which they have a right. The teacher, therefore, must see to it that punishments and reprimands are understood by the child. She must give opportunity to each child to experience joy and success in his work.

Swearing, when it has not become an habitual language response, is often an outlet for an emotional stress. The complaining type of individual, the grouchy man, is often letting off steam by means of his complaints, because he has been blocked in the direct expression of his emotions. Alcohol and drugs very often act as outlets for suppressed tendencies to reaction. When a superior blames us for something in our teaching, we vent our indignation upon the helpless children in the class. We blame them for what is obviously our fault. When we cut our finger because of our unskillful handling of the knife, we throw the unoffending knife away. When we miss our drive at golf, we hurl the innocent club to the ground.

Introversion. — Instead of responding directly to the things in our immediate environment, our search for an outlet for repressed modes of response may take another direction. We may turn in upon ourselves and find satisfactions in certain less explicit modes of response. The offending remark is made by the teacher and the child sits down, apparently calm and undisturbed. But if we could read his thoughts, we might be astounded at what is happening to the teacher there. He is making up a story in which he is being justified and the teacher properly and thoroughly humiliated. Defeated, unsuccessful and despised in school, the child may take refuge in his day-dreams, where without disagreeable effort, he marches forth resplendent in glory, conquering all his enemies and doing marvellous and noble deeds. The petty slights and wrongs, the frequent unfairnesses of the world, "the oppressor's wrong, the proud man's contumely," all disappear as by magic in our day-dreams. Released from external necessities, from the toil and exertion of overcoming obstacles, we attain our desires in imagination. If the child has no companions or unsatisfactory companions, he very often creates imaginary ones and plays with them and holds lengthy conversations with them. Discouraged or disgruntled with our position, we may become Napoleons and Alexanders in a "cleaner, greener land."

Desirable and Undesirable Outlets. — Outlets of some sort are necessary. Our original modes of response are very powerful and all through life they play their part, ready to burst forth in all their primitive intensity in new situations to which we have not learned to adjust ourselves. Many of the excesses common in war-time may be explained by the newness of situations to which men have not acquired modes of adjustment, and also to the lack of many of the customary restraining factors present in a normal peace-time environment. Introversion and extroversion are simply two types of adjustment. In both these types we find desirable and undesirable outlets. In the extroversion type of response obviously such outlets as drink, drugs, burglary and the like are undesirable. In the introversion type day-dreaming may not be undesirable unless carried to excess. A pronounced tendency to introversion has been picturesquely called a "flight from reality." If this flight from reality keeps us out of touch with our surroundings too long and if it paralyzes our active life, it becomes undesirable. If our day-dreams never spur us to action, if they tend to become completely satisfying in themselves, then we become more and more shut-in, more and more independent of the external world. The extreme case of this type is the form of insanity known as Dementia Praecox, where the patient may for days be totally oblivious to his surroundings, but, nevertheless, living an active imaginary life. Other abnormal forms of day-dreaming lead to delusions of grandeur and power.

Sublimation. — The question then remains as to what are the desirable outlets for the tendencies that must necessarily be repressed in any form of civilization. They may be either of the extroversion or introversion type and in most cases these types will be mixed. In general the desirable form of outlet will be such as to lead the individual to a mode of response that will be not only harmless, but will contribute something of positive value to himself or society. Finding solace in a pipe of tobacco, various kinds of swearing, a moderate amount of day-dreaming — all these are very common and harmless or neutral ways of finding an outlet. They may not be sufficient for the more dominant instinctive urges, such as the sex urge, and they certainly do not turn these tendencies to action into constructive channels. If a reprimand makes us grit our teeth and try hard to do better, if failure makes us attack the task again more doggedly, if our day-dreams result in ambitions which we actually plan and work to realize, then we may say that we have found desirable outlets for our instinctive modes of response. Failure as well as success may be a potent stimulus to further action. We must see to it that children's failures do actually lead to renewed attempts at mastery. Day-dreams may result in pictures of what we should like to be and act as a real spur to ambition. If the day-dream results in a concrete plan, in a poem or in a story, the outlet that our emotions have thus found is a desirable one. This turning of our instinctive urges into desirable channels has been called sublimation. It is a refinement, elevation or purification of the original mode of response. When the original tendency to rage or anger at some specific injustice done to an individual, say a case of child labor, makes a man devote his life to bettering the condition of children, to help pass laws for their protection, we may say that a sublimation of the instinctive tendencies to rage, anger and fighting has taken place. When a teacher rebels at the drudgery and poverty of her lot, and, instead of complaining about the unfairness of her superintendent or board of education, converts such energy into improving her own professional training or in working for the general betterment of the profession, we have the same process at work.

To be sure the necessity for outlets for the original tendencies of children have been known by teachers in all ages. Stern and repressive systems of education have paid less attention to them than more liberal and enlightened ones. A wise control of the sex life of the adolescent is not to be found in the strict segregation of the sexes and the treating as obscene all reference to sex. The modern educator attempts to give an opportunity for the sublimation of these tendencies by means of societies, dances, games and the like, where the two sexes may have adequate opportunities to meet and form the right attitudes towards each other. If the teacher is consciously aware of the ways in which our original tendencies function, it should be possible to arrange better opportunities for desirable outlets.

As one psychologist (Watts, 4) dramatically expresses it: "The human soul is a battlefield upon which the irrational impulses of the personality strive with the rational and ethical interests of the personality for the supremacy. The educator must justify his vocation by effectually assisting the latter forces to establish such a mastery that the conquered systems of desire will function not as rebels but as willing citizens, in the little kingdom of the mental life."

Rationalization. — Our original tendencies are among the strongest drives to action that exist. They form the underlying motives for much of our behavior. All through our lives, they are insistent. Even although man prides himself on being the rational animal *par excellence,* nevertheless, much of his behavior is directly motivated by his original tendencies. At the same time we are taught continually to act according to reason. We try to be reasonable beings. We search continually for the reasons for all happenings. We are not content until we can assign a reason for most things, and so it comes about that we try to find reasons for all of our own conduct. This necessity for finding reasons for his conduct is instilled in the child from an early age. He is constantly being asked, "Why did you do that?" The "why" is insistent and he is given no peace, until a plausible reason is given by him or suggested to him.

Now, as a matter of fact, many of our acts are the direct consequence of original tendencies and we frequently do not ourselves know the motive that has been at work, or, knowing it vaguely, we are not willing to delve back to this real motive and bring it forth into light. We are ashamed of many of these real motives and would rather not recognize them at all. They frequently do not agree with our own ideals, with the picture of our self as we should like to have it. The necessity for explaining our conduct either to others or to ourselves makes us, therefore, find some plausible reason for our conduct. This process of finding a reason for our behavior, which is not the real reason, is called rationalization. We rationalize or make reasonable our conduct, either because we do not know the real reason, or because we do not wish to acknowledge the real reason. We do not rationalize when we attempt to find the real mainspring of our action, and, finding it, bring it clearly before us, or, not finding it, acknowledge the fact that we do not know the reason.

Rationalization is seen in its simplest form in the excuse-making of the child. The child does not want to eat his oatmeal, and so it becomes "too hot" or "too cold" or "too lumpy." He does not want to do his arithmetic and so his pencil is "too long" or "too short" or his "pen won't write" or he "must do his spelling first." The student does not want to study and so he decides he "must call up that friend" or he "must take more exercise and have a walk." The father wants to live out of the city near the golf course and so he buys that suburban house, so that the family "may get more fresh air," and the children "have more space to play."

Particularly in the realm of moral conduct do we find many rationalizations for deviations from a strict moral code. The Prohibition Amendment has caused countless rationalizations among people who will not admit that they are law breakers. Hart (1) says very aptly, "Thus it is a familiar fact that people of otherwise irreproachable honesty will swindle the government or a railway company with untroubled equanimity. If they are taxed with the incongruity between their principles and their conduct, a varied crop of rationalizations will be immediately produced. They will point out that a company is not the same thing as an individual, that nobody really loses anything, that the fares or taxes are so inequitable that it is justifiable to evade them, and so on. The distinction between the real and apparent causes of mental processes is well illustrated in the advice given to the newly created judge, 'Give your decision, it will probably be right. But do not give your reasons, they will almost certainly be wrong.'"

As the child grows older we notice how the crude and unabashed "I want," springing from the original tendencies of his nature, gradually changes to "I should like this because" or "if I had or did this, I could. . . . " Now a development out of the crude statement of our wants or the direct satisfaction of our wants without regard to others, is of course necessary. We all rationalize, some more and some less. We must be careful, however, that the practice should not go too far. We must be on the watch, lest we build up a fictitious personality assigning only such reasons as this personality will allow, and at the same time satisfying our crude wants and desires under the protection of the ingenious rationalizations which we devise. The danger is that we shall become less and less able to face reality. We shall permit all our crudest and undesirable wants to find satisfaction under the guise of rational conduct, and at the same time imagine ourselves as strong, virtuous and rational beings. We must, so far as possible, therefore, face the real reasons for our own actions and attempt in like manner to help the child do so. This does not mean incessant lecturing and "reasoning" with the child. From much of the lecturing and reasoning done by the adult, the child obtains splendid suggestions for his own rationalizations.

The Complex. — When a great number of stimuli, which in other people lead to varied reactions and ideas, tend in a particular individual to arouse a given set of reactions or call up a given constellation of ideas, we say that the individual in question has a complex. Put in another way, a complex has been described as a system of emotionally toned ideas. When a boy says he is "crazy about radio," when numerous stimuli, unconnected with radio as far as the average person is concerned, all tend to make him think and talk about radio and when he spends all his money and time on radio, we say he has a "radio complex."

Complexes may be of all sorts, narrow or wide, temporary or relatively permanent. The fanatical reformer may be said to have a complex with reference to the particular reform which he is incessantly urging upon everyone he meets. The enthusiastic scientist may have a complex centering around a particular scientific hypothesis which he has propounded. In all these cases certain dominant ideas seem to take possession of the mind. The individual is set to respond in a certain particular direction to the innumerable stimuli crowding upon him. The individual may imagine that he is thinking and reasoning logically, when he is really selecting only those stimuli that fit in with the dominant trend of thought. In this way he becomes illogical, for he is selecting his evidence, and refusing to consider anything that may contradict his beliefs. The anti-vaccinationist feels so strongly that vaccination is wrong that he cannot think about the subject without getting emotionally stirred up. He cannot reason logically about it, because he steadfastly refuses to consider all the evidence. He thinks he is logical and rational, but arguments of the sort that would convince him in other matters are utterly wasted upon him in this particular case.

An "inferiority complex" is a term that has been applied to a general feeling of inferiority to others. Thoughts and expectations of failure are continually present. The individual with this complex is always willing to let others lead. Each new task is begun with the expectation of failure. In all plans and hopes for the future, the difficulties and dangers and chances of failure loom up more clearly than the thoughts of success. Similarly, we can speak of a "superiority complex," which makes the individual regard himself as vastly better than those around him, even although his actions belie his thoughts.

In several forms of insanity we find complexes more clearly present and more potent in their influence on behavior than is ordinarily the case with sane people. Every stimulus may set the patient off on his particular complex. For years he may have been relating everything to this one set of ideas, so that now it forms his entire universe. Complexes may be harmful or harmless, according as they dominate the individual's life. Any one set of thoughts that tends to crowd out all others will make for a badly balanced personality. The teacher must be on his guard against fostering a superiority or inferiority complex in any child. A certain amount of success is necessary for every child, as well as a certain amount of failure.

Freudianism. — A great many of the concepts and ideas we have been using so far in this chapter have been directly or indirectly introduced into our subject by Freud's system of psychology. Freud started with an interest in abnormal phenomena and gradually worked out a scheme of psychology always from the point of view of its direct value in the treatment of mental disease. This system of psychology has become so popular that a brief description of it may be valuable here. Some of its concepts are so valuable that they have found a permanent place in the academic psychology of the present time. On the other hand the whole system has by no means been accepted and is not likely to be so accepted by psychologists in general. References to Freud and his theories are, however, becoming so common in educational literature, that it is well for the teacher to have some understanding of the main scheme of Freud's theory.

Briefly, then, we may say that Freud postulates the existence of original tendencies in man, as do most present-day psychologists. He differs, however, from the majority of psychologists in maintaining that the supreme instinctive tendency is the sex instinct. Indeed, as he develops his system it would seem as if no other instinct had much motive power, except the sex instinct. He uses this term "sex" in a very broad sense, and would seem to include under it the affection of the parent for the young, and of the young for the parent and also that general desire to be with others of the same species which we have called gregariousness.

Now the varied manifestations of the sex instinct, or desires or wishes as Freud calls them, continually seek expression. A great many such expressions are permitted in our present civilization, but a great many are not. Hence a great many of our desires or wishes must be repressed. When they are repressed, they go down into the subconscious. They do not cease to exist but remain as forces in the subconscious. Here they seethe around seeking a way out into consciousness. Barring their way stands the censor, in Freud's picturesque language, who will not permit them to rise into consciousness in their crude natural forms, because as such they will offend and shock the individual. The more cultured we are, the higher are the standards of our censor. Hence the crude naked desires must clothe themselves in varied disguises in order to find an exit. This they do and so they emerge in various ways. In dreams they come to us with the least disguise, because when we sleep the censor is not so vigilant as when we are awake. Our dreams, therefore, are important as indicators of the repressed wishes and desires seething in our subconsciousness. If the desires assume a better disguise they emerge as driving forces in our lives and with the normal person they fit in or are sublimated and hence the idea of sublimation which we have already discussed. If sublimation does not take place we have the abnormal adjustment, and all the symptoms of the neurotic, hysterical or psychopathic individual are to be conceived of as modifications of repressed wishes.

Psycho-analysis. — If we wish to cure the patient of the mental disturbance, we must analyze his mental condition. We must go back from the present symptom to the initial wish or desire. Make clear to him what the real undisguised wish or desire was in the first place. Once he understands this and is allowed to give the wish free expression in some desirable way, the nervous difficulty will disappear. The analysis leads frequently back to early infancy and the Freudians emphasize the importance of this period of life for the beginning of abnormal tendencies.

This brief statement of Freudianism as a system of psychology and psycho-analysis as a method of treatment for psychical disorders must suffice in this place. There is no reason for an extended treatment in an educational psychology. Like all brief treatments of a complicated subject it fails to do justice to the real contribution of Freud. However, enough is given to enable the reader to follow the main lines of thought and also to indicate why the psychologist is not ready to accept Freud's theory. The complete dominance of the sex instinct is doubtful. Other instincts are important motives to action, particularly in childhood. It seems misleading and needlessly confusing to merge all childish loves for parents and siblings and friends into one sex instinct. Granted that the sex instinct is one of those which needs great modification throughout life, there is yet no need for a subconsciousness within which repressed desires remain as active forces. All this with the picturesque censor is utterly alien to the teachings of science. And there is no need for such hypotheses. The modifications of instinctive activities by means of habit are adequate to explain the building up of abnormal modes of behavior. And, lastly, although psycho-analysis may claim many cures, yet so also may Christian Science, Mesmerism, Our Lady of Lourdes, patent medicines, electro-therapy, blood-letting, homeopathy, and the King's Touch. The superior person may scorn the ignorant peasant plodding along to the confessional, while he himself is going to the consulting room of the psycho-analyst.

In place of analysis we must think of training. The hysteric, the psychopathic, is to be thought of as an individual whose reactions are undesirable and who may by proper training be taught to act and think normally. Belief in the possibility of relief is necessary and we may, therefore, think of all the above-mentioned methods of cure as being or having been potent means for training with some patients at some times, and yet they remain mere makeshifts which cover our ignorance of the real causes of the trouble. In its attempt to find causes, the method of psycho-analysis is to be commended.

Methods of Experimentation. — Various methods have been tried in order to study more objectively our original tendencies. We shall give a brief description of some of the more important of these and thereby, show the very different lines of approach which have been tried.

I. *Physiological Methods.* — The stirred-up feeling that we call an emotion is due to very decided physiological changes in the body, particularly in the glands. Exciting stimuli increase or decrease the normal secretions of these glands. The measurement of such increase or decrease is generally not possible in human beings, but much useful experimental work with animals has been accomplished. Stimuli causing the disturbances that we call fear or anger diminish very markedly the gastric secretions of the stomach and tend to check the churning movements of the stomach that pass the food along. Individuals under the influence of fear and rage are not able to digest their food.

Exciting stimuli also increase the secretion of the adrenal glands. The secretion, adrenalin, releases the supply of sugar stored in the liver. Testing for sugar in the blood is, therefore, used as a means of detecting the emotional effect of a stimulus. So far, this method has not been used very extensively, but it may contribute valuable information for psychology.

Methods of experimentation more directly applicable to the human being deal with the changes in breathing, in the pulse and in blood pressure. Emotional disturbances cause measurable changes in all these respects, but the interpretation of such changes is difficult. The psycho-galvanic reflex measures the changes in the resistance of the body to an electric current. Sudden changes in amount of resistance seem to be symptomatic of emotional disturbance.

II. *Dreams.* — This method is largely used by the Freudians. A study of an individual's dreams may help to reveal those instinctive tendencies and emotions which are more apt to be inhibited during his waking consciousness. The Freudians have worked out an elaborate system for the interpretation of the symbolism which occurs in dreams. To the general psychologist their methods appear very unscientific and unreliable. To Freud many dreams show the fulfillment of wishes. The desires that we cannot attain in our waking life are fulfilled in our dream life. "The dream substitutes itself for action." Thus the starving man dreams of a gorgeous banquet; the smoker who has stopped smoking dreams that he is enjoying a fine cigar; the lover dreams of his beloved. The dreams of small children seem to be frequently of this wish-fulfillment type. These straight-forward wish-dreams seem easy of explanation, but the interpretation of more complicated dreams is difficult and has not been satisfactorily accomplished. However much we may disagree with the wild speculations of the Freudians, we must admit that they have at least presented a challenge to psychology in insisting upon a psychological study of dreams.

III. *Word Reaction Methods.* — These methods all depend upon giving the subject a stimulus in words and asking him to react to it in words. They are usually called Association Tests, and can be divided into two large groups (a) Free Association and (b) Controlled Association.

A. *Free Association.* — Here we have three types: (1) the method of continuous association in sentences. Here the subject is just asked to relate his thoughts, the experimenter prompting and encouraging when necessary. This method is used largely in psycho-analysis. The starting point may be any complaint of the patient; it may be a dream or it may be some incident that the patient has previously related to the experimenter. Whatever the stimulus is the subject is asked to repeat any thoughts that occur to him on the presentation of the stimulus and to continue for a more or less indefinite period.

(2) Continuous series of words. Here the subject is given one word as a stimulus and is asked to repeat or write down all the words that occur to him for a given period of time. If the experimenter starts with the word "dog," the subject might respond by saying, "cat, house, tree, bird, wood, builder, noise," and so on. Here not only the words themselves but their number and sequence are of importance. A form of this test occurs in the Binet

Scale, where the number of words a subject can repeat in three minutes is used. As a written three minute test it has been given to children of various ages, and we note a great increase in the number of words as the child grows older, from about 20 at age 8 to about double that at age 15.

(3) Single word response. Here the subject reacts to the stimulus word by a single word response. This is generally done orally, the experimenter calling out the word and taking the time which elapses between the stimulus and the response. Any list of words may be used, but a very common list to employ is a series of one hundred words standardized by Kent and Rosanoff (2). It is standardized in the sense that the frequencies of all the different responses which occurred when the test was given to thousands of people have been listed. Hence we may tell how frequent or common any specific response word is. The authors of the test point out the difference between normal and insane subjects. Where the normal subjects give on the average 92 per cent common responses and only 7 per cent individual responses, the insane on the other hand give only about 71 per cent common responses and about 27 per cent individual responses. The insane person is more likely to differ in his thoughts and ideas from the average person around him and the association test brings out this difference by the greater number of unusual response words. The more eccentric a person is, the more likely is he to show a high percentage of unusual or individual responses. In no sense, however, must this test be taken to be diagnostic of insanity. A high percentage of individual reactions may occur in perfectly sane individuals, and also there are many insane people who show a high percentage of common reactions. Children below the age of eleven give fewer common and a greater number of individual responses than adults. Above age eleven the percentage of common and individual

responses is much the same as among adults. Evidently the younger children have not yet entered into that community of ideas which is common to individuals brought up in the same type of environment. Habits of thought are as yet more fluid and have not become as fixed as with adults.

B. *Controlled Association.* — In this type of test the subject is called upon to respond by a word of a certain sort. Only a limited number of words are deemed correct responses. Thus, if the stimulus word is "good," and the subject has been instructed to give the first "opposite" that occurs to him, he will obviously try to select one of the limited number of words which may be considered the opposite of good, such as "bad, evil, wicked, rotten, harmful," and the like. His reactions are "controlled" to this extent, in contradistinction to the giving of any word in the free association test.

There are many different kinds of controlled association tests. Some of the common types are: opposites (noisy—quiet); part-whole (leaf—tree); genus-species (taste—sweet), but it is possible to use many other relationships. Some of these tests, notably the finding of opposites, have proved very valuable in intelligence testing.

An Association Experiment. — Let us try one of these association tests so that we may become more familiar with the procedure of the test, and so that from our own results we may appreciate better the meaning and significance of association tests in general. The reader may do the test by himself, or else the test may be taken by the class as a whole with the instructor of the class acting as experimenter. If it is to be done as a class experiment, do not read any further than the end of this paragraph. Stop there and skip to page 58, continuing with the paragraph headed "The Uses of Association Tests." If you are going to do the test yourself now, provide yourself with a sheet of paper and a pencil. In any case, be sure you do not read Appendix I at the end of this book. Appendix I tells you how to score and gives you the results of the test. If you read Appendix I before starting on the test as described below, it will be useless for you to take the test. Your results will be worthless and have no significance whatever. Do not read or glance at Appendix I until you have taken the test.

Directions for the Test. — When all students have been provided with a sheet of paper, tell them to write the numbers from 1 to 18 vertically down the left hand side of the paper, so that they may write a word after each number. Then say, "You will be asked to close your eyes and then to think of something that belongs to a general class, for instance, musical instrument, amusement, algebraic symbol. If I should ask you to think of "a musical instrument," the first thing of that kind that might occur to you might be "piano" or "violin" or "cornet" or "flute" and so on. When the class of thing to be thought of is announced, notice what particular thing of that kind comes first to your mind. Do not search for anything else, but at once open your eyes and write down the particular thing you thought of. Be sure to write down the FIRST one that occurs to you. Please pay no attention to what your neighbor is writing. Does everyone understand? Get ready.

1. Number 1. Close your eyes. Write down the name of a color—the name of a color.
   (To the instructor—spell "color" and any succeeding stimulus word, if it is obvious that some of the students do not understand the word. Speak clearly and deliberately. Allow time for all to write — ordinarily a few seconds is sufficient. Proceed immediately tc the next word.)
2. Number 2. Close your eyes. Write down the name of an article of furniture — article of furniture. (Pause.)
3. Number 3. Close your eyes. Write down the name of a flower — the name of a flower. (Pause.)

4. Number 4. Close your eyes. Write down any letter of the alphabet — letter of the alphabet. (Pause.)
5. Number 5. Close your eyes. Write down the name of a metal — name of a metal. (Spell metal.) (Pause.)
6. Number 6. Close your eyes. Write down the name of an historic personage — an historic personage. (Pause.)
7. Number 7. Close your eyes. Write down a part of speech — part of speech. (Pause.)
8. Number 8. Close your eyes. Write down the name of a geometrical figure — geometrical figure. (Pause.)
9. Number 9. Close your eyes. Write down any verb — any verb. (Spell) v-e-r-b — verb. (Pause.)
10. Number 10. Close your eyes. Write down the name of a tool — the name of a tool. (Pause.)
11. Number 11. Close your eyes. Write down an article of food — article of food. (Pause.)
12. Number 12. Close your eyes. Write down a part of the body — part of the body. (Pause.)
13. Number 13. Close your eyes. Write down the name of a day of the week — day of the week. (Pause.)
14. Number 14. Close your eyes. Write down the name of a room in a house — room in a house. (Pause.)
15. Number 15. Close your eyes. Write down the name of an animal — an animal. (Pause.)
16. Number 16. Close your eyes. Write down the name of a book — name of a book. (Pause.)
17. Number 17. Close your eyes. Write down a girl's name — a girl's name. (Pause.)
18. Number 18. Close your eyes. Write down the name of a country — a country. (Pause.)

Now turn to Appendix I at the end of the book and score the papers according to directions. The experiment is described fully by Pintner (3).

The Uses of Association Tests. — As we have seen from our experience with the Community of Ideas Test, which we have just taken, there is a great deal of similarity in our word responses to given stimulus words. Where there are many marked deviations from the common response, there must evidently be a reason for these deviations. Emotional disturbances are likely to cause such deviations. If we study a number of the deviations from the common responses on a long free association test, like the Kent-Rosanoff Test, we may be able to find some connection between them. They may all refer to some general idea or activity. In other words they may lead us to the discovery of a "complex," a group of emotionally toned ideas. If something is weighing on our mind, if we are worried about something, if we have recently had an experience that has caused us considerable emotional up-set, it is very likely that the stimulus words of an association test will call forth response words that are more or less directly connected with the worry or exciting experience. In this way two general uses of the association type of test have arisen.

Psychiatrical. — In attempting to understand the mental difficulties of a patient, the association test has been employed. The responses of the patient may cluster around one or two central ideas or complexes and in this way a beginning of the study of the difficulty may be made. The physician will prepare his list of words so as to tap the complex which he suspects may be present. He will take into account the length of time taken to react to each word.

A very brief example given by Jung and Peterson may illustrate the method. Below are given the stimulus words, the responses given by the patient and the reaction times:

Stimulus Word	*Reaction Word*	*Time in Seconds*
1. Head	Hair	1.4
2. Green	Meadow	1.6
3. *Water*	*Deep*	*5.0*
4. Stick	Knife	1.6
5. Long	Table	1.2
6. *Ship*	*Sink*	*3.4*

Stimulus Word	*Reaction Word*	*Time in Seconds*
7. Ask	Answer	1.6
8. Wool	Knit	1.6
9. Spiteful	Friendly	1.4
10. *Lake*	*Water*	*4.0*
11. Sick	Well	1.8
12. Ink	Black	1.2
13. *Swim*	*Can Swim*	*3.8*

Stimulus words 3, 6, 10 and 13 all result in lengthened reaction times, and 13 also shows a peculiar response. All these words center about water and pointed to a suicide complex which dominated the patient.

Legal. — Some attempt has been made to use the association test for the detection of guilty knowledge, on the theory that the guilty person will have some group of emotionally toned ideas surrounding the crime he has committed. The experimenter, knowing the nature of the crime, inserts significant words among the standard list of stimulus words. The significant words are words which might have some connection with the crime under investigation, but they would in general appear to have little connection with such crime by the innocent individual. The stimulus words are then presented and the time taken to react to each word carefully noted. Next day, without previous warning, the association test is repeated.

The experimenter has four clues to guilty knowledge. The first clue consists of significant reactions to words showing a knowledge of the crime. Such naïve reactions would be rarely given by the cautious criminal. The second clue is the time taken to respond. If certain words call up as a first response a word connected with the crime, the guilty individual may suppress such a response and go on in his mind seeking a second or third word which is apparently innocent. Now this suppression and search for other words takes time, when time is measured in fifths or tenths of a second. All words that take appreciably longer than the average time of response, are, therefore, brought together and scrutinized to see if they point to a knowledge of the crime. The third aid to the experimenter is the repetition of the list, for if during the first test the individual has suppressed the first response and gone on seeking another apparently innocent response, he has thereby disobeyed the directions for the test and he will not in general be able to remember until the next test what particular responses he made in these cases. The great drawback about lying is that we cannot remember the lies we have told.

The fourth clue is the number of stereotyped responses, that is, similar responses or responses of a definite type, indicating that the subject is not following the instructions accurately, but has prepared himself beforehand for the sort of thing he will say. He is not giving just the first word that comes into his mind and hence suspicion is cast on him.

This is a brief description of the association test as applied to the detection of falsehood or crime. It depends upon the emotional state of the guilty man. Obviously where the thief is pursuing his work in the spirit of an occupation or profession, his day's work may not arouse any emotional response, just as the ordinary routine of the honest man does not, and in this case the association test will be of no use. Again the association test can only tell us where a suspicion of guilt lies, it cannot prove guilt. As a matter of fact, it has not been used in the practical investigation of crime, but the possibilities of such practical use are obvious.

The association test is a good example of one of the means used in experimental work with our emotions and instinctive tendencies. We shall see in the next chapter how it can be modified for use in the measurement of non-intellectual traits. As we have described it in this chapter, it is of no particular practical value in the school. It is interesting, as I have suggested, to give the Community of Ideas Test to school children, and one may learn things from so doing, but the test itself is of no immediate practical value.

Summary

1. This chapter supplements the preceding chapter by describing what happens when the original tendencies are at work.
2. Conflict frequently arises between the original tendencies and the demands of the environment.
3. When a conflict arises an outlet is sought, either of the extroversion or introversion type.
4. Extroversion finds an outlet in a reaction to the environment external to the individual.
5. Introversion is the type of outlet that drives the individual back upon himself.
6. Outlets, either of the extroversion or introversion type, may be desirable or undesirable, and it is the business of education and of mental hygiene in particular to help the individual find desirable outlets.
7. Sublimation is the refining of undesirable original modes of response, and is the name used for the finding of desirable outlets.
8. Rationalization is the process whereby we invent, consciously or unconsciously, motives or reasons for our actions, so that our actions may conform to a rational mode of life.
9. The complex is a constellation or group of ideas clustering around some central idea. This group of ideas is emotionally toned and dominates much of our behavior.
10. Freudianism is a system of psychology built up by Freud to explain the behavior of man, and particularly to throw light on his abnormal behavior.
11. Psychoanalysis is a method of treatment for certain forms of abnormal behavior.
12. Physiological methods of studying the emotions have increased our knowledge of the bodily changes during emotion.

13. The analysis of dreams is supposed by Freudians to give an insight into many hidden motives of behavior.
14. Word reaction methods are used to gain information pertaining to our emotions and instinctive tendencies.
15. Such methods are used to unearth complexes and may be of help to the psychiatrist. They may also help in the discovery of guilty knowledge but they cannot be commonly used in legal procedure.

Review

True-False Statements

As a review mark the following statements true or false:

1. The best method of adjustment to the thwarting of our fundamental impulsions is the substitution of some wholesome but vigorous activity.
2. A timid person, lacking in ability to make adequate social adjustments, frequently takes refuge in day dreams as a substitute for action.
3. One is always aware of the fact when he is rationalizing his conduct.
4. Most cases of mental conflict are due to the sublimation of instinctive tendencies.
5. The free association test may be of use in the discovery of complexes.
6. By the term "censor," Freud means the various conventions and taboos which restrain the free expression of many original tendencies.
7. When we honestly endeavor to find the real motive for our acts, the process is called rationalization.
8. Substitutions for thwarted impulses never lead to great achievements.
9. Repressed wishes in the "subconscious" are probably in the interior of the brain.
10. Parents and teachers should always insist upon a child giving a reason for his misbehavior.
11. The best suggestion for mental health is the intelligent facing of reality.
12. A group of emotionally toned ideas is known as a complex.
13. The subjects on which we are "touchy" probably indicate complexes.
14. Association tests give positive proof of the guilt of criminals.
15. Sublimation is the turning of a repressed original tendency into a socially desirable and useful channel.
16. Extroversion and introversion, being unnatural modes of reaction, are invariably undesirable.
17. In the functioning of original tendencies, conflicts arise between our original tendencies and our environment.

18. A child's imaginary companion is an example of introversion.
19. The Freudian subconscious seems to be a region into which we thrust our repressed desires.
20. Dream interpretation is for Freud a method for the discovery of complexes.
21. People who give few common responses in an association test are undoubtedly insane.
22. Some of us who "hate mathematics" may have been conditioned against it by having a teacher whom we disliked.
23. An intelligent child should always be expected to give a good reason for his acts.
24. Self-pity is a good example of extroversion.
25. As all original tendencies have evolved in the slow and long struggle for existence, they are valuable and should be thwarted as little as possible.
26. Extreme outward behavior of one sort may be an attempt to hide an inward state which is directly opposite.
27. Indicate the probable outlet involved in each of the following reactions by putting the appropriate number before each. (1) Extroversion; (2) Rationalization; (3) Sublimation; (4) Introversion.

( ) a. A slow pupil imagining himself to be some day a great leader.

( ) b. A pupil making excuses to his parents for a low mark in Arithmetic.

( ) c. A pupil muttering in a class.

( ) d. A boy, holding a grudge against his class-mate, working off the grudge in a vigorous game of tennis.

( ) e. A pupil playing the rôle of a bully on the playground.

( ) f. A man scrubbing floors in an asylum claims he is a millionaire.

( ) g. Why struggle for mastery and leadership? — the supreme virtues are inherent in meekness.

( ) h. A woman having lost her husband and child in a terrible accident throws herself wholeheartedly into her work as head of an orphanage.

( ) i. A man low in the scale of authority in his place of work rules his wife and children with an iron rod.

( ) j. People extraordinarily competent along one line must be deficient in other lines.

( ) k. A man of insignificant physique walks with dignified gait and manner and talks in a loud masterful voice.

( ) l. A pretty girl has no brains. (Spoken by a homely girl.)

28. Underline the best response:
When a suppressed original tendency seeks a desirable outlet, it is called: (1) introversion; (2) extroversion; (3) sublimation; (4) rationalization.

29. Underline the best response:
A group of emotionally toned ideas is called: (1) repression; (2) a complex; (3) the libido; (4) psycho-analysis.

Advised Readings

Morgan, J. J. B.: *The Psychology of the Unadjusted School Child*, New York, 1924.

Wells, F. L.: *Mental Adjustments*, New York, 1917.

Gates, A. I.: *Psychology for Students of Education*, New York, 1923, Chapters 6, 7, 8 and 9.

References

1. Hart, B.: *The Psychology of Insanity*, Cambridge University Press, 1912, Chapter 5.
2. Kent, G. H. and Rosanoff, A. J.: "A Study of Association in Insanity," *American Journal of Insanity*, Volume 67, 1910.
3. Pintner, R.: "Community of Ideas," *Psychological Review*, Volume 25, No. 5, Sept., 1918, pages 402–410.
4. Watts, F.: *Abnormal Psychology and Education*, New York, 1924, page 146.

CHAPTER IV

THE MEASUREMENT OF NON-INTELLECTUAL TRAITS

Personality. — When we talk about the personality of an individual we evidently mean those traits, characteristics or qualities of an individual which are found in him. Psychologically his personality would be a catalog of all the different responses made to all the different situations that confront him. It would be a complete picture of the man's reactions, and under the term reactions we include all the things he says and all the thoughts he thinks. It is, of course, impossible to make such a catalog. All that we can ever hope to do is to note some of the common tendencies to respond to certain types of situations. In ordinary life and in literature what we do when we describe the personality of a man is to note those characteristics which are particularly strong in him, those that stand out unequivocally, those that mark him off from his fellows.

Ordinary Descriptions. — Our description of a man might be as follows: He is a medium-sized man, rather stockily built, with broad shoulders and large head. He is not physically strong and does not enjoy good health. He has made more than an average success in his profession and is keenly interested in his professional work. He does not, however, enjoy a very widespread reputation among his professional co-workers, although some rate him very highly. He is clear-cut and decisive in his professional thinking and shows courage in defending his own particular views and opinions. In other matters he appears rather timid and is decidedly timid where physical danger is concerned. He enjoys social gatherings and has a great number of friends and acquaintances. He enjoys indoor games, but rarely plays any outdoor ones. He is poor in administrative matters, hesitating to come to a final decision and is likely to let such things slide. He is punctual in attendance to his duties, and shows high ability in handling his work. He is widely read and possesses a vast amount of well-coordinated knowledge. He is slow to anger and keeps his emotions well under control at all times.

This is an unorganized random description of an individual such as we find frequently in literature, but it brings out very well the heterogeneous mass of characteristics that spring up when we talk of the personality or character of an individual.

It is an interesting class exercise to ask students to jot down four or five phrases descriptive of well-known men, say Roosevelt and Wilson, and then have the class discuss the possibility of grouping these descriptive phrases into broad categories. Taking the replies of one class at random, we have such phrases as these:

Roosevelt	*Wilson*
altruistic	writer of interesting letters
big stick	man of letters
rough and ready	high idealism
aggressive	an introvert
physically weak when young	courageous
writer of books	died broken-hearted
fat	thin
very informal	independent
typical blusterer	brilliant mind
extrovert	leader
overcame obstacles	aristocratic
great knowledge of birds	a scholar
big physically	liked to travel
an opportunist	very self-centered
quick to learn	single track mind

These are samples of phrases written by different students and it is not our purpose here to judge the correctness of their opinions. The lists are merely indicative of the sort of thing that results from general descriptions of people. If now we try to group these phrases we get words like "fat," "thin" and the like referring to the physical appearance of the men. Then there are certain phrases like "brilliant mind," "quick to learn," indicative of the intelligence of the individual. Then there are phrases such as "man of letters," "scholar," "great knowledge of birds," pointing to the knowledge or skill possessed by the individual. Lastly and by far the most numerous we find a heterogeneous group of phrases, such as "introvert," "extrovert," "leader," "informal," "died broken-hearted," "aggressive," "liked to travel," and so on, having reference to the emotional make-up of the individual, his moral and personal qualities, his likes and interests.

Classification of Traits. — We may, therefore, make a first rough tentative classification as follows:

1. Statements with reference to physical qualities or physical appearance.
2. Statements with reference to ability to do things, to think, to learn.
3. Statements with reference to knowledge, skills, or things learned.
4. A heterogeneous group of statements referring to emotions, such as anger, moral qualities, such as conscientiousness, and such qualities as punctuality, courage, and also such things as likes, dislikes, interests and ambitions.

With the first set of statements we have, as psychologists, little or nothing to do, except in so far as they may be significant for conduct or the explanation of conduct. The second set of statements refer to what the psychologist has set apart as intelligence, the innate ability of the individual to learn, to be modified or changed along certain lines, and a discussion of the measurement of intelligence as these qualities have come to be called, is given in another chapter. The third set of statements deals with the individual's acquired knowledge or skill; it is concerned with what he has learned, what modifications in an intellectual sense have been acquired, whether such acquirements are abstract, like ability to read Virgil in the original, or concrete, like ability to drive an automobile or paint a house. The last or heterogeneous group of statements has not yet found one word to describe it. It is with the measurement of these qualities that this chapter is to deal, and hence we have called them non-intellectual traits to distinguish them from the intellectual traits and the different kinds of knowledge and skill which are now so well measured by psychology.

This group of non-intellectual traits contains all the emotions and all those traits commonly designated character qualities. It includes also personal traits that are not generally dignified by the name of character traits. Moral qualities in a narrow sense are also included, and so also are such things as likes, dislikes, interests, ambitions, ideals, hopes and aspirations. This jumble of traits has not as yet been clearly analyzed by psychology. Many a theoretical attempt at classification has been made, but no one such classification has been accepted. Undoubtedly a clearer picture will emerge when once we are able to measure such qualities. Measurement will bring with it sharpness of definition and clarity in classification.

Dependence on Original Tendencies. — All these traits help to make up what is commonly known as the character of an individual. Psychologically we may think of character as the total mass of conduct tendencies of the individual. What a man is likely to do in various situations is the index of his character. And what he is likely to do, can of course only be surmised from a study of what he is doing now and what he has done in the past. When we study the conduct of an individual we find a set of habits of action, more or less loosely organized. These modes of reaction are based upon his original tendencies, but they have been conditioned by his environment. We cannot get rid of the habits or modifications that have been built upon the original tendencies, but the varying strengths of the original tendencies have undoubtedly influenced the modifications that have arisen, and hence the diversity of characters that occurs among individuals brought up in the same environment. The characters of brothers and sisters may differ greatly in spite of the similarity of their environments. The problem in the measurement of character qualities, then, is to arrange test situations by means of which the measurement of any trait or traits may be made. In arranging these we are not concerned as to whether the reactions called forth are instinctive or acquired. All that we want is a measurement of the responses which the individual now makes. If there is a close agreement or correlation between the responses made on the test and a certain type of reaction in everyday life, then we say that the test is valid, that it is an index of the type of reaction in question.

Measurement of Such Traits. — The growth of measurement of character traits has been very slow and we are only now in the experimental stage of such work. There is little that can be of immediate practical value to the teacher, but a survey of what is being done will be of interest. There is a tremendous research activity in this field at the present time, and we cannot here refer to all the work that is being done. All that we can do is to pick out a few samples for purposes of illustration. In making such a survey we will begin first with some descriptions or inventories of the conduct reactions of young children. Here we shall see character qualities in the making, without any definite attempt to measure or evaluate them, merely to note their appearance in the young child. Secondly, we shall pass on to the rating scale which is a device for estimating how much of a given trait an individual possesses. This is merely an attempt at a more accurate description. We watch an individual or we depend upon our previous knowledge of him, and then say that he possesses so much of this trait and so much of that, or that he is better than this other individual in this or that trait. Thirdly, we arrive at the test proper in which the individual makes certain responses, which are evaluated and scored. The score is the measure of the amount of the trait in question.

Descriptions or Inventories. — The descriptions we shall mention here are not the usual literary descriptions which we discussed at the beginning of this chapter. Such descriptions are of comparatively little interest to the psychologist. They depend too much upon the person making the descriptions, upon his subjective judgment and his individual standards, and they are also made in very general terms. There are, however, some very careful and detailed descriptions of the reactions of infants which are helpful in a beginning of any study of character. By studying a number of infants, we can arrive at a common or general tendency. Gesell (2) has made such summaries for young children. In giving such summaries of the reactions of young children he emphasizes the individual differences which are present. "Even the four-months-old child appears to have his individualities and idiosyncracies. Indeed almost every infant is likely to have some trait of character, some trick, habit, achievement, or peculiarity which expresses this individuality and which occasionally may be symptomatic of a defect, weakness, or excellency."

Some of the items in Gesell's summary for a four-month-old baby are: Notices large objects; shows selective interest in animated face; not much affected by strange persons, new scenes or solitude; laughs aloud, etc. At six months we note: expresses recognition of familiar persons; may show consciousness of strangers; enjoys presence and playfulness of persons, etc. At twelve months we note: cooperates while being dressed; inhibits simple acts on command; imitates simple acts like scribble and spoon-rattle. These are only a few items out of many at definite stages of growth from birth to five years. These general conduct tendencies may serve in a way as character tests for the infant. They show types of response that are in general developed at specific stages of growth. These summaries are not exactly tests but they serve as standards in our examination of infants.

Ratings. — Many rating scales for all sorts of characteristics have been attempted. Some depend on ranking individuals in order of merit. Others depend on the comparison of the individual to be rated with individuals well-known to the rater, which individuals have already been arranged in an order of merit. This type of rating scale has been called the man-to-man rating scale and it was largely used in the army for the rating of officers. Another type of rating scale, frequently called the graphic rating scale, makes use of the distance between the opposites of a given trait. Terman (7) has made use of this type of rating scale for comparing gifted and control children. Twenty-five separate traits are involved, but there can be just as many or few as the investigator desires. Each trait is represented on a long line as in Figure 4.

The rater is asked to compare the child he is rating with the average of the same age. He has to make a cross anywhere on the line where he thinks the child belongs. These positions on the line are then turned into

Trait 4. Self-Confidence

Extreme self-confidence and self-reliance. Always relies on own judgment. Courts responsibilities.	Decidedly self-confident	Rather self-confident	Average for age	Rather self-distrustful	Decidedly self, distrustful	Extreme lack of self-confidence. Distrusts own judgment. Afraid of responsibilities.

Trait 20. Truthfulness

Extraordinarily truthful, honest and frank. Never misleads or misrepresents, however great the temptation.	Decidedly more truthful than average	Rather more truthful than average	Average for age	Rather less truthful than average	Decidedly less truthful than average	Extreme tendency to lying, deceitfulness and evasiveness. Lies for the slightest advantage.

FIGURE 4. — Graphic Rating Scale. (From Terman's *Genetic Studies of Genius*. Stanford University Press.)

scores from 1 to 13, 1 being the highest possible rating, 7 the average and 13 the lowest. With this method one can score just as finely as he likes, but he should be careful not to score more minutely than differences in the trait which can be appreciated by the average rater. A finer scale than 1 to 13 would be of doubtful value. In Terman's investigation both parents and teachers rated many children on many traits, and there is a slight tendency for them to agree in their judgments of children.

Undoubtedly such a graphic rating system is one of the best methods devised, but, nevertheless, it suffers as do all rating schemes from the subjectivity of the performance. Our judgments of individuals are at best haphazard and influenced by all sorts of prejudices and peculiarities. The precise meaning attached to a trait varies from rater to rater. What I have in mind when I think of self-confidence is not exactly the same as what you have in mind. Furthermore, we are influenced by our likes and dislikes. If we like a person, we are more likely to rate him in general high and, conversely, if we dislike him. We naturally think of individuals as "wholes" rather than as conglomerations of separate traits, and so we judge one person to be "good and fine" and another to be "poor and weak." This general impression casts a "halo" over the individual and all our specific judgments of his traits are influenced by the "halo" of the general impression which surrounds him. All this makes the rating of individuals difficult and unsatisfactory. In lieu of anything better, we use rating scales, but they will disappear at once as soon as adequate objective tests of character qualities appear. In actual school work rating scales may be used to good advantage by the teacher to make him study and think over the personal characteristics of each child in the class, but the teacher should be warned not to place much emphasis upon the results of such ratings.

Tests. — Tests for the measurement of personal and moral traits are the outcome of very recent research, so recent, indeed, that the results and value of such tests as exist are much in doubt. It would be useless here to attempt a survey of all the different sorts of things that have been tried. We shall merely choose a few sample tests of different aspects of character and try to describe what they are and how they propose to measure character qualities.

Emotions. — People are conceived as differing in the amount of emotional response that will be caused by the same situation and also in respect to the number of situations that will evoke an emotional response. The Woodworth Personal Data Sheet is an attempt to measure the general emotionality or the emotional stability of the individual. A modification of this questionnaire has been prepared for the use of children by Mathews (4). It consists of a number of questions to which the child responds by underlining "yes" or "no." These questions may be roughly classified as dealing with:

1. Fears, worries and the like.
2. Physical symptoms, pains, weariness.
3. Unhappiness, unsocial and anti-social moods.
4. Dreams, phantasies, sleep disturbances.

Some of the actual questions are:

Do you like to play with other children?
Would you rather play by yourself alone?
Do people find fault with you much?
Are you afraid of the dark?
Do you dream of robbers?
Do you ever stutter or stammer in your speech?

There are 75 to 100 questions of this kind. No significance is attached to any one answer, and, of course, the psychologist knows that he cannot expect the truth to all questions from every child. The total number of "unfavorable" responses does, however, seem to be significant. The greater the number of such responses, the more likely is the child to be emotionally unstable. We cannot as yet say that there are specific standards or norms for this test, nor can it be regarded as a sure diagnosis of emotional instability. It seems to have decided value as a survey of large groups of children. Those who make particularly high scores may then be given a detailed examination by the psychologist and psychiatrist. This test is a good sample of one of the attempts to penetrate into the difficult field of testing the emotional make-up of the individual.

Character Traits. — Among the large number of traits under this heading, we find psychologists attempting to measure such specific traits as aggressiveness, suggestibility, confidence, sociability, decisiveness or speed of decision, self-assurance and the like. Many different methods of attack have been tried. Let us take as a sample the tests for suggestibility by Otis (5). A group test for use with children has been devised. Each child is given a test sheet and the examiner gives instructions as to what is to be done with each item, just as if it were a test of following directions, indeed many items are pure directions items and the suggestibility items are scattered at random throughout. The suggestibility items run all the way from things that are very obvious, as with a picture that might well be the head of a horse where the child is told to underline "yes" if it looks like a race horse, to questions that are absurd, such as, "Draw a line under one of these words (pine-trees, maples), if it answers this question, "Do acorns come from pine-trees or from maples?" or to such direct suggestions as, "Here are names of different fruits. Draw a line under the name of the one you like best. Most people like plums best."

This test has been given to many children and it would seem to show that ability to resist suggestion increases as the child grows older. The adult shows a much greater resistance than do children in general. It would seem also that the more intelligent child is less open to suggestibility than the less intelligent one. This trait of resistance to suggestion is not a simple clear-cut trait but is obviously closely akin to such qualities as self-assertiveness, self-confidence, good judgment, self-reliance and auto-criticism.

Moral Traits. — Certain tests have been made to test more specifically the moral and ethical qualities. There have been many workers and many tests. The recent work of Hartshorne and May (3) will be our best example. Their work on the measurement of deceit, or the tendency to deceive, or the trait of honesty is the most extensive piece of research in character measurement that we possess up to the present time.

They have constructed a battery of tests, both individual and group, to measure the tendencies to deceit as exhibited by school children. Some of their methods are:

Duplicating Technique. — An intelligence or educational achievement test is given, taken to the office and a record of the child's performance made. Then the unscored paper is returned to the child, together with a scoring key and he is asked to score his own paper. The number of corrections and additions made is a measure of his honesty.

Improbable Achievement Technique. — Here we have tests where the child can easily fake the answer and in so doing show his dishonesty. If a difficult puzzle is given with a time limit within which an expert cannot solve it and if a child solves it easily within this time limit, it is very probable that he has not been honest. The peg solitaire puzzle was one used in this way. Weight discrimination of seven small pill boxes, marked in order on the bottom, but differing so slightly in weight that no one could arrange them in order, except by chance, was another test used with this technique. Here also belong the various "peeping" tests, where mazes are to be traced or numbers put in circles with the eyes closed—feats impossible of accomplishment if honestly done. Only those who "peep" can do them.

Double Testing Technique. — Here the children are tested twice; once under strict supervision, and the second time under lax conditions where copying answers from the key or changing answers can be accomplished. For this purpose duplicate forms of reliable tests are used and the difference between forms under no cheating conditions must be known for the population to be tested. If now, differences more than three times the standard deviation of the non-cheating differences are found, then the fact of cheating is probable and the amount of cheating can be measured. Many different types of tests were used with this technique, such as arithmetic, information and other educational tests, as well as cancellation, substitution and similar simple speed tests.

In addition to these several methods used in school, they have tested honesty in different situations:—

Honesty in Home Work. — One form of a word knowledge test is given to the pupils to take home with instructions not to get help from any person or dictionary. Another form of the same test is then given in school under strict supervision. Amount of deception is measured as described above.

Honesty in Athletic Contests. — A contest was arranged for the school. Four tests were given: (1) dynamometer or strength of grip; (2) spirometer or lung capacity; (3) chinning or pulling up on the horizontal bar; (4) the standing broad jump. In all cases practice trials were given, followed by other trials which the pupils recorded themselves. The deception score is the difference between the highest of the practice trials and the highest of the self-recorded trials.

Honesty in Parlor Games. — Here certain parlor games, such as Pinning the Tail to the Donkey, Bean Relay, The Mystery Man and the like were used in carefully controlled situations to measure the children's honesty.

In addition there were methods to measure the stealing and the lying types of deception.

This brief summary of the types of tests used must suffice here. The group tests for school children in the classroom were used very extensively by Hartshorne and May. As to the validity of such tests, they say, "In general and test for test, we are able to measure deception almost as consistently and with almost as little error as we are able to measure intelligence." One such test of deception is of little value, but "if we use ten tests of classroom deception, however, we can safely predict what a subject will do on the average whenever ten similar situations are presented." If we want to measure the average deceptive behavior of large groups of children, a much smaller number than ten tests will suffice.

These tests were given by the investigators to many school populations and the results have been presented in great detail. No other moral trait has been so thoroughly measured and reported upon up to the present time. We may very briefly summarize the authors' chief conclusions: Older pupils are slightly more deceptive than younger children. Sex seems to make no difference. Honesty is positively related to intelligence; in groups of children of approximately the same age, those of higher levels of intelligence deceive definitely less than those of lower levels. Children who show symptoms of emotional instability are more likely to deceive than those with fewer such symptoms. Deceit is definitely associated with the economic level of the home and also with the cultural background. Children of North European ancestry show less tendency to deceive than those of South European ancestry. Colored children deceive more than most of the white groups. Little difference was found between children belonging to the three main religious groups, Catholics, Jews, and Protestants. Deception runs in families to about the same extent as eye color, length of forearm and other inherited structures. In most tests there are no grade differences. Children over-age for their grade cheat less than those under-age. Those getting high marks cheat less than those getting low marks. Pupils rated high in deportment cheat less than those rated low. There is considerable resemblance in amount of cheating between friends in the same classroom. There is less cheating where the relations between teacher and pupil are free and cordial than where this is not the case. Progressive schools show less cheating than conventional schools. Children who attend Sunday-School regularly cheat in day school about the same as those who rarely or never attend Sunday-School. Deceit is not a unified trait. Most children will deceive in certain situations and not in others. The motives for cheating, lying and stealing are complex and inhere for the most part in the general situations themselves. The most common motive for cheating on classroom exercises is the desire to do well.

Interests. — In addition to personal characteristics and moral traits, individuals differ markedly in their interests. What we are interested in influences our total outlook upon life, our activities and in a general sense our future. Older people obviously differ markedly in their interests. Children, as all teachers know, differ greatly in this respect also, and it would seem as if such differences in interests began at a very early age and were somewhat influenced by the inherited tendencies and capacities of the child. However that may be, the emergence of definite interest trends is well known and their significance for educational and, later in life, for vocational guidance is great. By the time a child comes to school and increasingly so during his school life, definite interests are present and it is important for education to measure them, if possible.

Many attempts have been made to tabulate children's interests by having the child check lists of things or indicate his preferences or indirectly show his interests by means of information tests. All of these schemes have given us some indication of the different interests of children, but they have been very unreliable and subjective. Recently an attempt has been made (Terman, 6) to measure interests more objectively by means of the association type of test such as we have described in the previous chapter. The test consists of a very carefully compiled list of 120 stimulus words. Not any word will do, as in the ordinary free association test, but each word must be equally adapted to provoke responses due to intellectual, social or activity interests, for these were the three general classes of interests which it was proposed to measure. The children were simply asked to write down in one word what the words shown them made them think of. All the response words given to these stimulus words were then given a score value for each of the three types of interest. These weights or values were determined by the responses given by a group of children who were well known by their teachers. The method is too complicated to describe in detail here. In this way, for example, the response "diamond" to the stimulus word "gem" is rated 20 for intellectual, 11 for social, and 15 for activity interest. If the stimulus word "grand" evokes the response "piano," then this is rated 8 for intellectual, 9 for social, and 11 for activity interests. Each response word receives a score for each of the three types of interest. A child, therefore, receives as a result of giving responses to all the stimulus words three total scores—one for each type of interest. This is as it should be, for no one is solely interested in one aspect of life. The three totals give a measure of the amount of interest in the three different directions, and a comparison of the three scores with each other shows the main or dominant interest. Thus a boy, who is described thus "would rather read than play—likes history—no leadership—dramatic ability—enjoys music," has a high score for intellectual interests, and only a moderate score for social and activity interests. A girl described as "musical—quiet and unassuming—extraordinary amount of physical energy" has a high activity interest score and average scores for social and intellectual interests. If this test proves as successful in general use as it has with the originator of the test, it would seem as if we are on the way to a very satisfactory and objective means for the measurement of children's interests.

General Temperament. — The most ambitious attempt to measure the total personality is the carefully constructed Downey Will Temperament Test (1). It is an attempt to give a picture of a great many personality traits, particularly those that are included under the term temperament. The temperament of an individual is presumed to show itself in various patterned forms of activity. These patterns are determined by (1) the amount of nervous energy at the disposal of the individual and (2) the tendency of such energy either to discharge itself immediately into motor areas or to find a round-about pathway of discharge. Temperament shows itself, then, in reactions of all sorts, in our speech, in our habits of walking, writing and doing all sorts of things. For test purposes walking or talking are difficult to control and measure and, therefore, writing has been chosen as the type of activity for use in a test. We must note carefully that this is not an attempt to judge character from handwriting, as in the pseudo-scientific study of graphology. Writing is merely used as a convenient type of response which will be indicative of the general habits of response of the individual.

The tests are constructed both for individual and group testing. The subject is asked to write his name in his usual style and speed, as rapidly as possible, as slowly as possible. He is asked to write other things in these three ways. He is asked to disguise his handwriting as much as possible, and to copy various styles of writing, sometimes as quickly as possible, sometimes at his own speed. He is made to write with his eyes open, his eyes closed, while counting and with other forms of distraction. There is also a memory test, a speed of decision test and a suggestibility test. In this way a number of personality traits are tested and each one is measured on a scale running from 0 to 10. Each trait is scored separately and a psychograph, or picture in the form of a graph, showing the strength or weakness of each trait is constructed. This is called the will-profile and is presumed to be a measure of the temperament of the individual. Some of the items of this profile are: speed of decision, self-confidence, finality of judgment, interest in detail, perseveration, motor impulsion and the like. A strong dominating personality shows high scores in most of the traits, and the reverse is true of a weak vacillating individual.

At present it is difficult to appraise this attempt to measure the general temperament. The reliability of the test is doubtful. The validity of the test is difficult to determine. In the hands of some workers it has shown very low correlations with estimates of character as made by other persons, but of course these character ratings are probably in themselves very unreliable. It has been tried with school children, but it is not suited for younger children. At present, therefore, we must regard it as in the experimental stage and not suited for practical use in schools. It is well worth consideration, however, as being one of the first attempts to measure objectively the total temperament of an individual.

Conclusion. — As a final comment to this chapter, let it be said again that mere samples of the vast amount of research work have been taken. Furthermore, we have not attempted to review all the work that has been done. For such a review, see Watson's book mentioned in the Advised Readings at the end of this chapter. The most promising type of work seems at present to be that done by Hartshorne and May. The general importance of all the work that has been done lies in the fact that we will in the future be able to measure those more subtle changes in reaction tendencies, subsumed under the general heading of character traits and moral qualities, which educators have always been attempting to improve. At the present time, however, there are no easily administered tests which the schoolman can use.

Summary

1. This chapter deals with the attempts of psychologists to measure non-intellectual traits. A few sample studies are cited. It is not an attempt to review all the research work that has been done in this field.
2. General descriptions of character are quite useless for scientific purposes, hence the need for tests.
3. Useful inventories of the reactions of infants are a beginning of character measurement at this early stage.
4. Rating scales are an attempt to make our judgments of character qualities more objective.
5. Rating scales, however, are still very subjective and their reliability is not great.
6. The Woodworth Personal Data Sheet is an attempt to measure the general emotional stability of the individual.
7. The Otis test for suggestibility is an example of an attempt to isolate and measure a specific character trait.
8. The best example of the measurement technique in this whole field is to be found in the work of Hartshorne and May with their battery of Deception Tests.
9. These tests have shown high reliability and validity in the measurement of deception for groups of school children.

10. These tests have been given to large populations and important relationships between the tendency to deceive and other factors have been discovered.
11. An attempt at the objective measurement of children's interests by means of the association test technique has been described.
12. An attempt to measure the general temperament of an individual has also been described.

Review

True-False Statements

As a review mark the following statements true or false:

1. In making a rating scale, it is essential that the scale be divided into minute divisions in order to care for individual differences.
2. Two of the Hartshorne-May classroom deception tests will give a good measure of a child's general tendency to cheat.
3. It is possible now to predict accurately whether or not a boy will become a delinquent.
4. The Downey Tests attempt to measure such traits as honesty and truthfulness by means of an analysis of handwriting.
5. The "halo" effect is eliminated by the man-to-man type of rating scale.
6. A sure sign of marked aggressiveness is to be able to look another in the eye without flinching.
7. To measure a person's honesty would require a very large battery of tests.
8. The Woodworth Personal Data Sheet attempts to measure such traits as perseverance and flexibility.
9. Temperament is supposed to show itself in various habitual forms of activity.
10. Association tests have been used in an attempt to discover the dominant interest trends of the child.
11. Honesty seems to be positively correlated with degree of intelligence.
12. Children who come from poor homes are just as likely to be honest in school work as children who come from wealthier homes.
13. A person's character can be very well determined by his handwriting.
14. The resemblance between siblings in deception is about the same in amount as the resemblance in eye-color.
15. Fill in the blanks:
    Three well-known types of rating scales are called:
    (1) ______________________________
    (2) ______________________________
    (3) ______________________________

16. Complete this sentence:
The character of a man can be thought of as the total number of his ———— ————.
17. Underline the best response:
There exist now fairly objective tests for the measurement of: generosity, sociability, initiative, honesty.
18. Complete this sentence:
The Woodworth-Mathews questionnaire was designed to measure ————————.

Advised Readings

Hartshorne, H., and May, M. A.: *Studies in Deceit*, New York, 1928. Introduction and Chapters III and XXIII.

Watson, G. B.: *Experimentation and Measurement in Religious Education*, Association Press, New York, 1927, Chapters III and IV.

References

1. Downey, J. E.: *The Will Temperament and Its Testing*, Yonkers, 1923.
2. Gesell, A.: *The Mental Growth of the Pre-School Child*, New York, 1925, Chapter 32.
3. Hartshorne, H., and May, M. A.: *Studies in Deceit*, New York, 1928.
4. Mathews, E.: "A Study of Emotional Stability in Children," *Journal of Delinquency*, Volume 8, 1923, 1–40.
5. Otis, M.: "A Study of Suggestibility of Children, *Archives of Psychology*, No. 70, 1924.
6. Terman, L. M.: *Genetic Studies of Genius*, Stanford University Press, 1925, Volume I, Chapters 16 and 18.

CHAPTER V

THE MEASUREMENT OF INTELLECTUAL TRAITS

Intelligence. — We saw in the last chapter in our general description of an individual that in addition to character and moral qualities, there are a certain number of traits that we may call intellectual traits. It is with these traits that we are now concerned, and we shall be solely concerned in this chapter with such traits from the point of view of ability to accomplish, rather than from the standpoint of what has been accomplished by the individual. A measurement of the knowledge acquired by the individual will come later in the second part of this book. Here we are concerned with the possibility of measuring his general ability to acquire knowledge. This general ability to learn has been called general intelligence. We shall for the time being accept this term without further description. Later in this chapter we shall attempt a more detailed description and definition.

Judgment of Intelligence. — This ability to learn is the most important trait of the child from the point of view of the educator. Most things taught in school have their intellectual aspect. A great many make their appeal solely to the intellect. We, therefore, are making use of a child's intelligence continually, and every teacher knows that children possess all degrees of intelligence. Some are slow in learning and some are quick; some learn easily and some only with great effort, and to others some things are forever barred. A teacher is constantly sizing up her pupils from this general standpoint of ability to learn. We are all doing this all through life, with reference to our friends, our colleagues, our competitors. How successful we are, it is difficult to tell. We know that we often make gross mistakes, as witness the frequent use of the expression, "I didn't think he had it in him," and the like. Nevertheless, a great many people have much faith in their ability to size up the intelligence of a person even after a very brief interview. We cannot here interview real persons. Let us, however, make our problem more concrete by making judgments of the intelligence of children from their photographs.

A Practical Experiment. — This experiment may be done either by the reader himself or it may be arranged as a class experiment by the instructor. In either case no reference to Appendix II should be made until after the experiment, otherwise the results will be worthless.

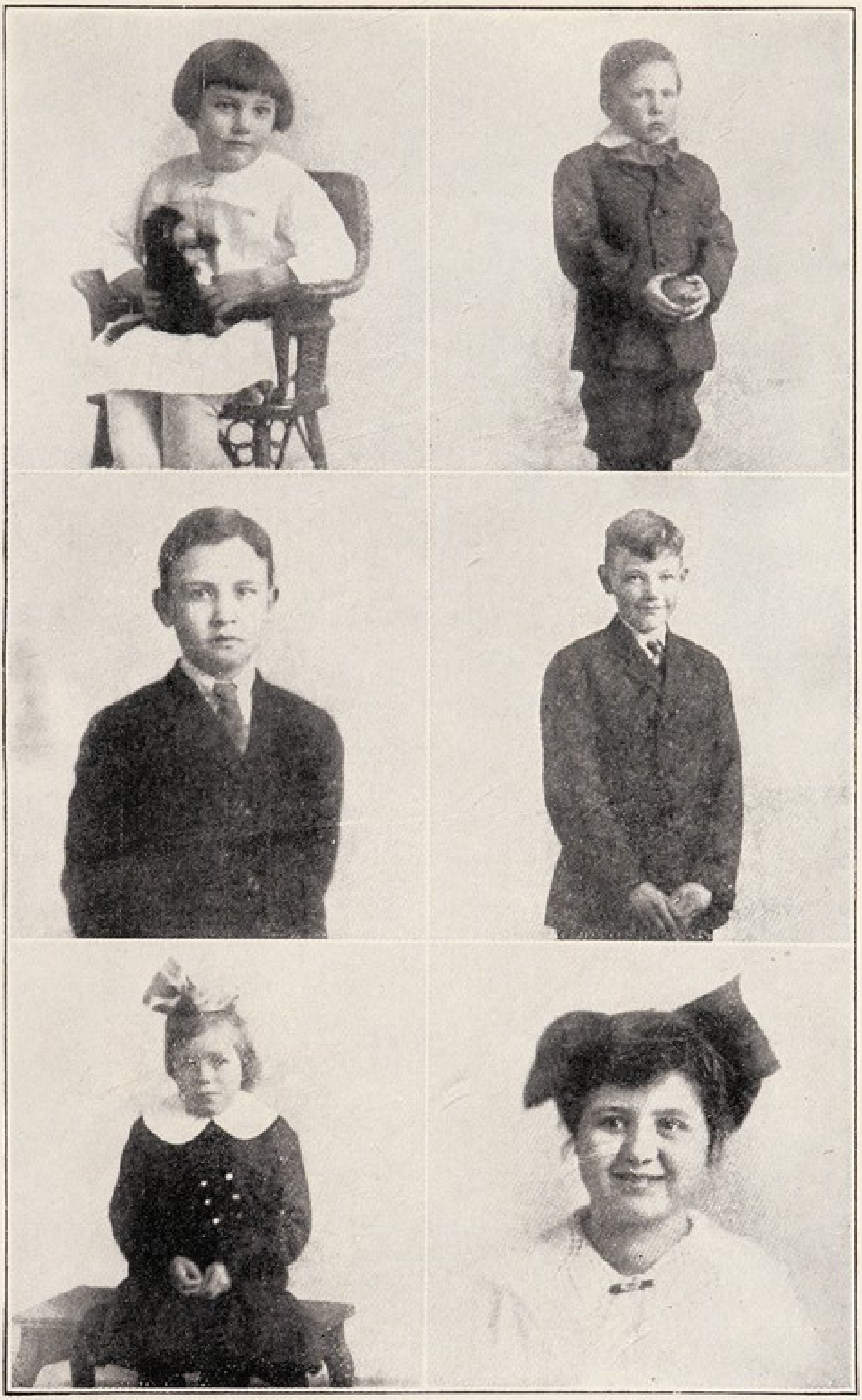
See that each member of the class is provided with a sheet of paper. Tell them to begin at the top left hand side and write as heading of the first column "child" and under this write down vertically, using one line for each letter, the letters "A" to "L" inclusive. Now tell them to draw vertical lines down their paper making five columns in all. These five columns are to be headed as follows: Column 1. "Child"; Column 2. "My Rank"; Column 3. "Test Rank"; Column 4. "Difference"; Column 5. "d^2". Now tell them to turn over the page where the photos of the children are printed, each child being designated by a letter. Say to the class, "Here are twelve photographs of children. You are to rank these children in the order of their intelligence. The children range in intelligence from very bright to very dull or stupid. Study them carefully. Decide which one you think is the brightest and put the number "1" opposite the letter for this child in Column 2, under "My Rank," on your paper. Now look for the next brightest child and record him as number "2," and so on until you reach the dullest child, who will be recorded as number "12." Allow

A B

E F

I J

FIGURE 5(a)



C D

G H

K L

FIGURE 5(b)

sufficient time for all to complete the ranking, but do not allow the students to discuss during the ranking and do not allow them to ponder over it too long. This is unnecessary and tedious for everyone concerned.

While the students are doing the ranking the instructor should write on the board the same kind of table for judging as the students have on their papers. He will use this as a sample of what the students are to do in working out a rank correlation. He should take the paper of one student who has completed the ranking and copy on the board the ranks in Column 2 under "My Rank."

At this point it is well for the instructor to go over the judgments quickly by calling for a show of hands. Ask who rated Photo A among the best three (i.e. ranks 1, 2 or 3) and jot the number of hands down on the blackboard, then ask who rated Photo A among the poorest 3 (i.e. ranks 10, 11 or 12). Do this rapidly for all the photographs. In a large class it will be found that someone has generally rated each photo either as among the three best or the three worst. Here the point can be made that people in general certainly differ tremendously among themselves as to which child is most intelligent and which is least intelligent, and that such judgments conflict radically with each other.

The instructor should now turn to Appendix II and write on the blackboard the ranks of the children according to the test. He should do this slowly having the class copy them on their papers. He should begin with rank 1 and comment on the intelligence ratings as he goes along. Give the class time and opportunity to note the differences between the test ranks and their own ranks and to make the appropriate reactions of surprise and amazement that are sure to follow, if the students have not studied Appendix II previously.

A Rank Correlation. — After the test ranks have been copied down by each student, tell them to subtract the "test rank" and "my rank" columns, paying no attention to signs. Demonstrate carefully with the example on the board. These differences show by how much they deviated from the true rank of each photo. Then in the column "d^2" write the squares of the values in the "difference" column. Add this last column and multiply by 6.

Now write down the formula for rank correlation, namely $\rho = \frac{1-6\Sigma d^2}{n\,(n^2-1)}$, and tell the students to copy this on their papers, explaining that they are calculating a coefficient of correlation according to this formula and that they have each their own particular value for $6\Sigma d^2$. Explain what Σ(sigma) and d^2 mean. Tell them that "n" is the number of cases and that $n(n^2-1)$ is equal to 1716 and will be the same for every student. Now solve the equation on the blackboard, explaining each step carefully and then have the students solve their own examples. If $6\Sigma d^2$ is greater than 1716, then a negative correlation will result. Be sure to explain how to handle such a case. Make a rough distribution of the coefficients of correlation on the board, or in a large class have two students prepare such a distribution, showing median and quartiles, and report to the class the next day. Compare this distribution with the table of results given in the appendix.

Value of Judgments of Intelligence. — Evidently our judgments of intelligence obtained from photographs are valueless (Pintner, 3). If any particular individual happens to have obtained a high coefficient, he should not be unduly proud of his achievement. If he were to repeat this test again in six months or so, after he has forgotten the ranking of these particular photographs, the chances are that his coefficient of correlation this time would be very low. If he tries another set of twelve pictures, the correlation is likely to be very different. In other words we have no evidence that any individual possesses the ability of estimating the intelligence of individuals from their photographs, although many people believe they can tell a lot about an individual from his photo. Hence the custom of having applicants for positions send a photograph. Apart from telling something as to the beauty or ugliness of the applicant, nothing can be discovered with reference to the intellectual or moral qualities.

Photos not a Fair Test. — Objection may be made to our experiment, and rightly so, that teachers do not judge the intelligence of the pupils from photographs but from the living subject. If we were to conduct an experiment with actual children standing in front of the class, the correlations might be a little higher. If the children were interviewed and asked questions for a minute or so each, the correlations would undoubtedly increase. If each of us could have these children as pupils in school for six months, then our ability to judge their intelligence would be very greatly increased. And finally, if we were to interview each child, using definite standard questions, we should approach the best measure of his intelligence, for in one sense, this is what an intelligence test is. The intelligence test presents the same definite situations to each child, and in accordance with his manner of response to these situations we judge his intelligence.

Tests are Standardized Interviews. — The intelligence test is a sort of standardized interview. It is standardized in the sense that we know how children of various degrees of intelligence are likely to respond to our questions. There is nothing mysterious or magical about it. It is simply doing in a better way, what we all do in daily life, what the teacher, principal, or employment manager is doing when conducting an interview, namely, inferring the amount of intelligence of the candidate from the quality and number of responses he makes to our questions or tests. The very important addition made by the psychologist is in the standardization. It tells him how to evaluate the different kinds of responses that may be made to a given situation. Let us now take a glance at the different kinds of tests which are used to estimate intelligence.

Different Types of Tests. — We may divide the numerous intelligence tests that are now available into different types. From the standpoint of the number of subjects examined at one time, we have the useful division into individual and group tests. We may also differentiate tests according to the content of the test material, whether verbal or non-verbal, or again as to the manner of giving directions for the test and making the appropriate responses, whether with use of language or without language (non-language). Other tests in which the subject is called upon to do things with his hands are called performance tests, to differentiate them from the tests which are largely or entirely oral or written. Again we might differentiate tests according to the age of the subject for which they are best adapted. There are tests suitable for the various ages, from the pre-school child up to the university student. Finally we may divide our measures into tests of general ability as contrasted with tests of special capacities. The tests of general abilities are what we think of as intelligence tests, whereas the tests of special capacities are restricted in scope, such as tests of musical talent, of artistic ability, of mathematical ability, and the like. There are as yet very few of such special capacity tests. Let us discuss some of the characteristics of these different types of tests.

Individual Tests. — The individual method of testing intelligence was the first method to arise and the most commonly used scale to-day, the Stanford Revision of the Binet-Simon Tests, is a revision of the original scale constructed by the French psychologists, Binet and Simon.

The Stanford Revision (Terman, 7) is applicable to American children brought up in the usual American environment which assumes attendance at school from about age six onwards. It consists of a miscellaneous group of about ninety tests arranged in groups of about six or more for most ages from age three to age fourteen with the addition of tests for two higher levels, called average and superior adult. These tests are given orally by the examiner and the child is scored right or wrong. Only so many tests are given as is necessary to determine the intelligence level of the child. The whole scale is not given to every individual. Beginning with a level well within the child's ability, the examiner proceeds as far as possible up the scale to an age level where all the tests are failed. Here are some samples of the tests at different age levels:

Age Three: Naming Familiar Objects. The examiner shows the child a key and asks, "What is this?", and similarly with a penny, a closed knife, a watch, a pencil. If three responses are correct the test is passed.

Age Four: Comprehension. The examiner asks the child (a) What must you do when you are sleepy? (b) What ought you to do when you are cold? (c) What ought you to do when you are hungry? Two correct responses are required. What is to be called "correct" and what "incorrect" has been standardized, and is therefore not determined by each individual examiner's own judgment.

Age Five: Comparison of Weights. Two weights, of 3 and 15 grams, are placed before the child about two or three inches apart. The examiner says, "You see these blocks. They look just alike, but one of them is heavy and one is light. Try them and tell me which one is heavier." Three trials with position of blocks reversed are given. Credit is allowed if two of three comparisons are correct.

Age Six: Missing Parts. A standard card is shown with faces having parts omitted. The examiner says, "There is something wrong with this face. It is not all there. Part of it is left out. Look carefully and tell me what part of the face is not there." Credit is allowed if three out of four are correct.

Age Seven: Description of Pictures. Certain standard pictures are shown the child, and he is asked, "What is this picture about? What is this a picture of?" Credit is allowed if two out of three pictures are described or interpreted. There are definite standards for enumeration, description and interpretation.

Age Eight: Finding Likenesses. The examiner says, "I am going to name two things which are alike in some way and I want you to tell me how they are alike. Wood and coal, in what way are they alike? In what way are an apple and a peach alike? In what way are iron and silver alike? In what way are a ship and an automobile alike?" Credit is given if two out of four are correct. Again what is to be considered correct has been standardized.

Age Nine: Four Digits Backward. The examiner says, "Listen carefully. I am going to read some numbers and I want you to say them backwards. For example, if I should say 5—1—4, you would say 4—1—5. Do you understand? Ready now; listen carefully, and be sure to say the numbers backward." (a) 6, 5, 2, 8 (b) 4, 9, 3, 7 (c) 8, 6, 2, 9. Credit is allowed if one of the sets is correct.

Age Ten: Naming Sixty Words. The examiner says, "Now I want to see how many different words you can name in three minutes. When I say ready, you must begin and name the words as fast as you can, and I will count them. Do you understand? Be sure to do your very best, and remember that just any words will do, like 'clouds,' 'dog,' 'chair,' 'happy'; ready; go ahead." If sixty words are given within the time limit, the test is passed.

Age Twelve: Dissected Sentences. The examiner points to the first group of words below and says, "Here is a sentence that has the words all mixed up, so that they don't make any sense. If the words were changed around in the right order they would make a good sentence. Look carefully and see if you can tell me how the sentence ought to read."

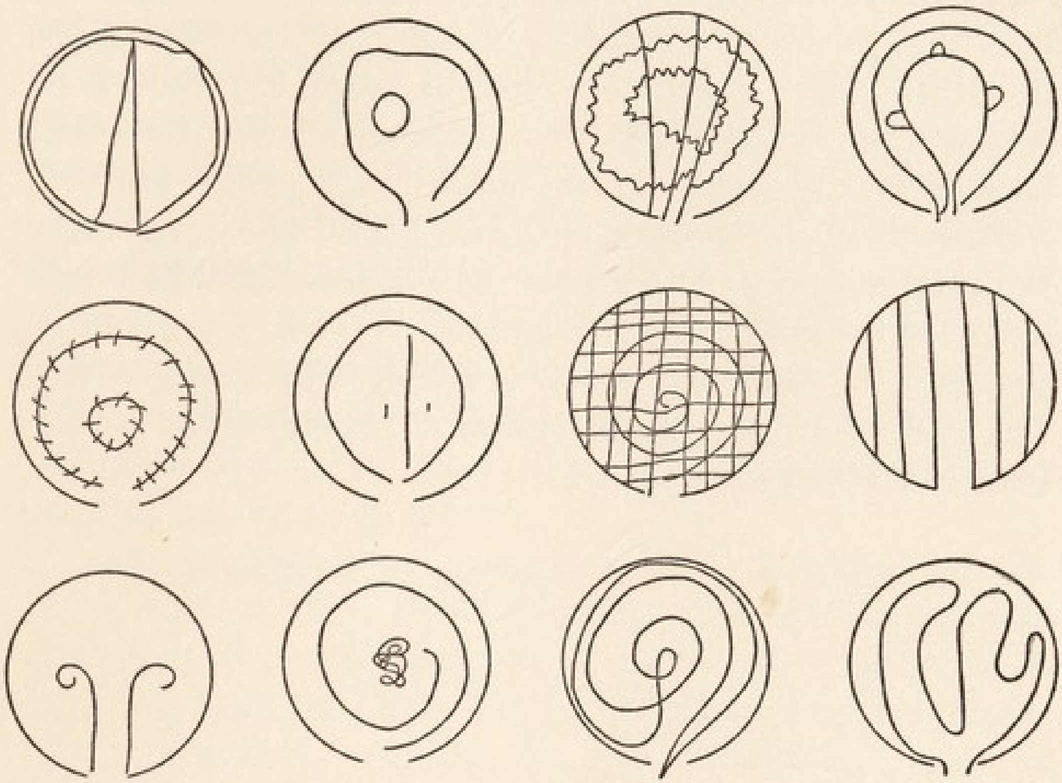


FIGURE 6(a). — Reactions to the Ball and Field Test of the Binet Scale. Responses credited at Year VIII.

1. FOR THE STARTED AN WE COUNTRY EARLY AT HOUR.
2. TO ASKED PAPER MY TEACHER CORRECT I MY.
3. A DEFENDS DOG GOOD HIS BRAVELY MASTER.

Credit is given if two out of three are correct, with a time limit of one minute for each sentence.

Standard Procedure. — The examples of the Binet tests given above show one test for each age level from three to twelve. There are three higher levels. The reader must remember that we have given only one out of at least six tests for each level. It is important to note further that the way in which each test is presented to the child has been carefully standardized. To deviate from the prescribed method of giving is to change the test, and to change the test means that we cannot use it in this scale because we do not know to which level it belongs. Furthermore the scoring of the responses has also been carefully standardized. Figures 6(a) and 6(b) show samples of reactions to the Ball and Field Test, where the child is told to trace out a path to show how he would hunt for a ball lost in a round field. Credit is given at year eight if the child more or less systematically covers the field. So, in a similar manner, standards for all the tests have been established.

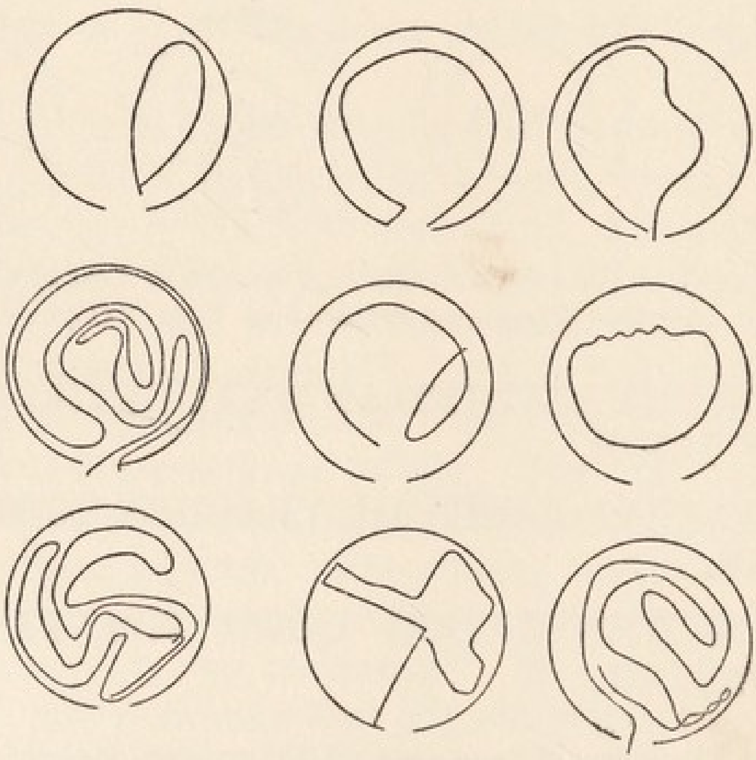


FIGURE 6(b). — Reactions to the Ball and Field Test of the Binet Scale. No credit at Year VIII.

Mental Age Rating. — From the examples of tests given, we see the miscellaneous character of the Binet Tests. The aim obviously is to obtain a sampling of many different kinds of reactions, so that the final result will be an index of all-round general ability. This final result is expressed in terms of mental age. The mental age of a child is an expression of that child's ability in terms of the average ability of children of like chronological age. If a child is said to have a mental age or M.A. of 8 years, we mean that he has accomplished on our tests as much as the average eight-year-old child accomplishes. All the tests have, therefore, been given to many children of different ages in order to discover the average performance at different age levels. It is this procedure that is known as standardization. The tests are valuable because they have been well standardized. It is only when tests have been standardized, that we are able to attach significance to a particular score or mental age made on a test.

The Intelligence Quotient. — Having obtained the M.A. on a test, we know the amount of intelligence possessed by the child, but we do not know how bright or dull the child is. This will be determined by comparing the child's M.A. with his C.A. or chronological age. If his M.A. is higher than his C.A., obviously he is more advanced than the average child. If his M.A. is less than his C.A., he is more retarded. This comparison of M.A. with C.A. is best made by computing the ratio of M.A. to C.A., and this ratio is called the Intelligence Quotient or I.Q. Intelligence Quotients from 90 to 110 show normal intelligence development, those below indicate dullness and those above indicate superior ability.

Other Binet Revisions. — In addition to the Stanford Revision of the Binet Test, there are several other revisions used in this country. Notable among such revisions is the Herring Revision of the Binet-Simon Scale. It tests the same sort of ability as the Stanford Revision does, and it forms a valuable alternative to the latter. If children have been frequently tested on the Stanford, they become familiar with the tests. In such cases the Herring Revision should be used. There are revisions of the Binet Tests adapted for use in almost every civilized country in the world, so valuable has the basic idea underlying the tests proved to be.

Performance Tests. — Other individual scales of intelligence, departing radically from the Binet-Simon type of test, have been constructed. The Pintner-Paterson Performance Scale is most commonly used. This scale instead of asking the children questions as in the Binet, gives them something to do or perform, hence the name. It consists of form boards of many types and the ability of the child to solve these different problems presented to him in concrete materials is measured.

In the Short Pintner-Paterson Performance Scale (4) there are ten tests as follows:

1. The Mare and Foal Test.
2. The Seguin Form Board.
3. The Five Figure Board.
4. The Two Figure Board.
5. The Casuist Form Board.
6. The Manikin Test.
7. The Feature Profile Test.
8. The Ship Test.
9. The Picture Completion Test.
10. The Cube Test.

Figures 7, 8, 9, and 10 show four of the above tests and from these pictures a good idea of the type of reaction demanded can be formed. No verbal directions are required.

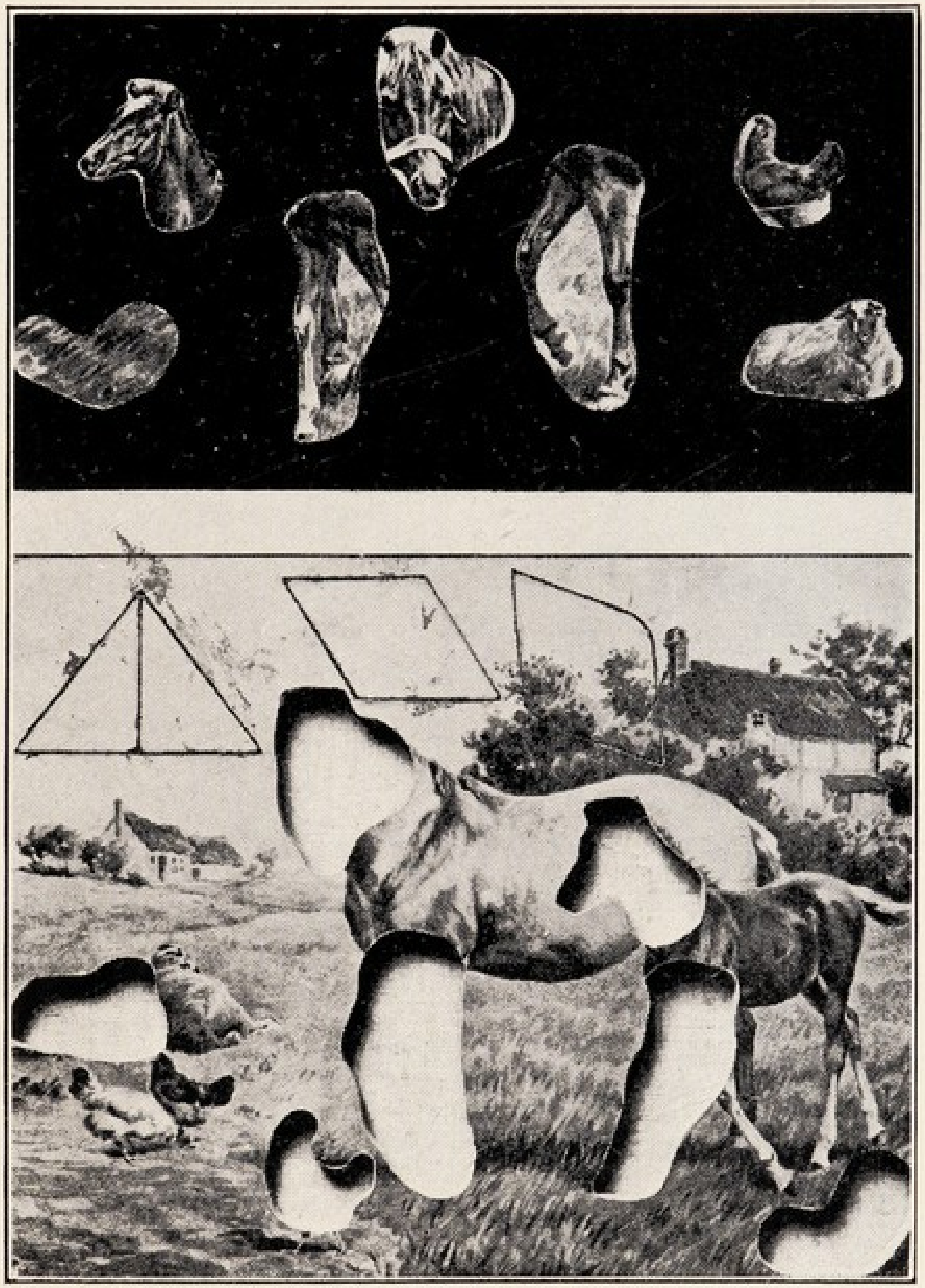


FIGURE 7. — The Mare and Foal Test. (From Pintner and Paterson's *A Scale of Performance Tests.*)

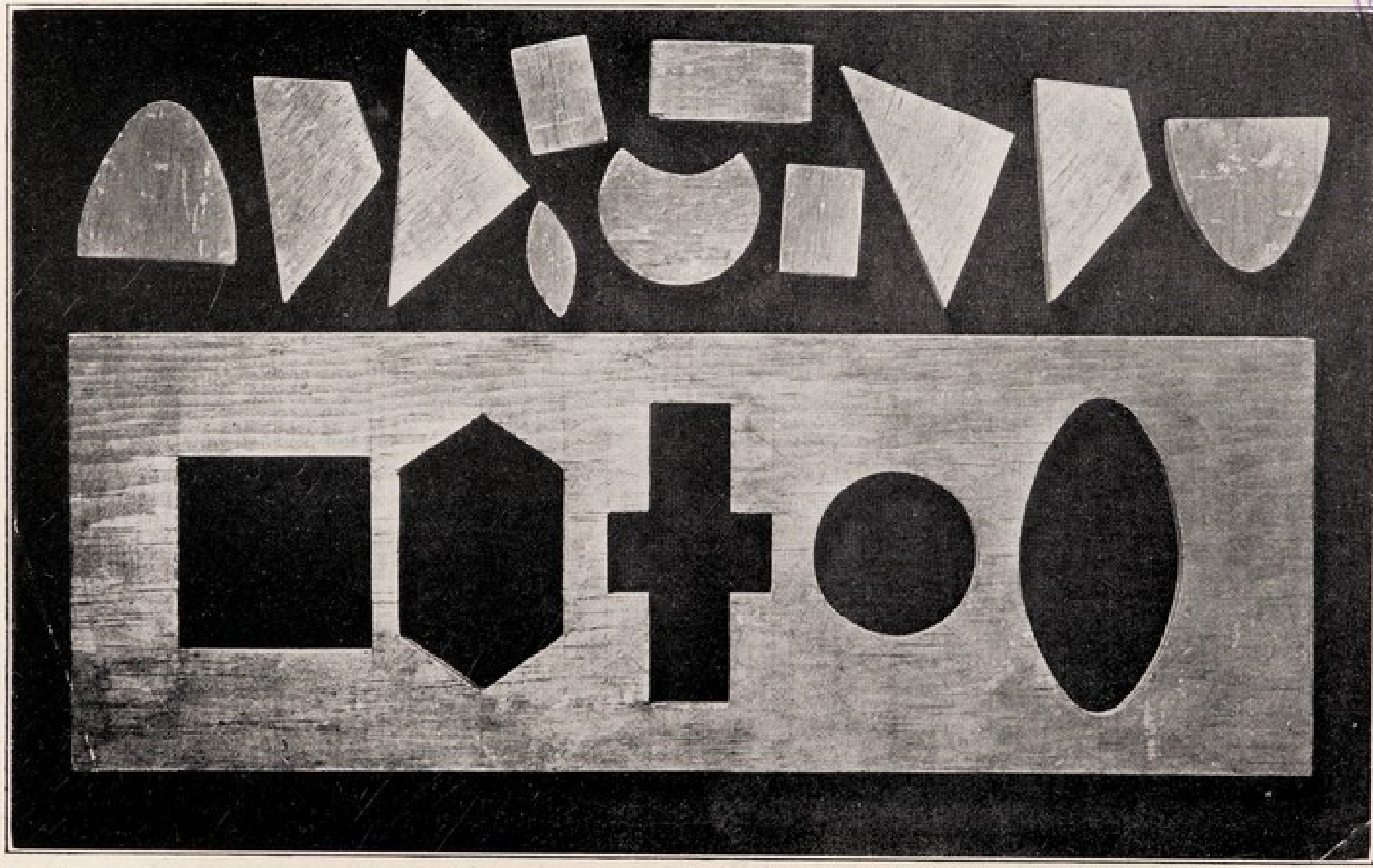


FIGURE 8. — The Five Figure Board Test. (From Pintner and Paterson's *A Scale of Performance Tests*.)

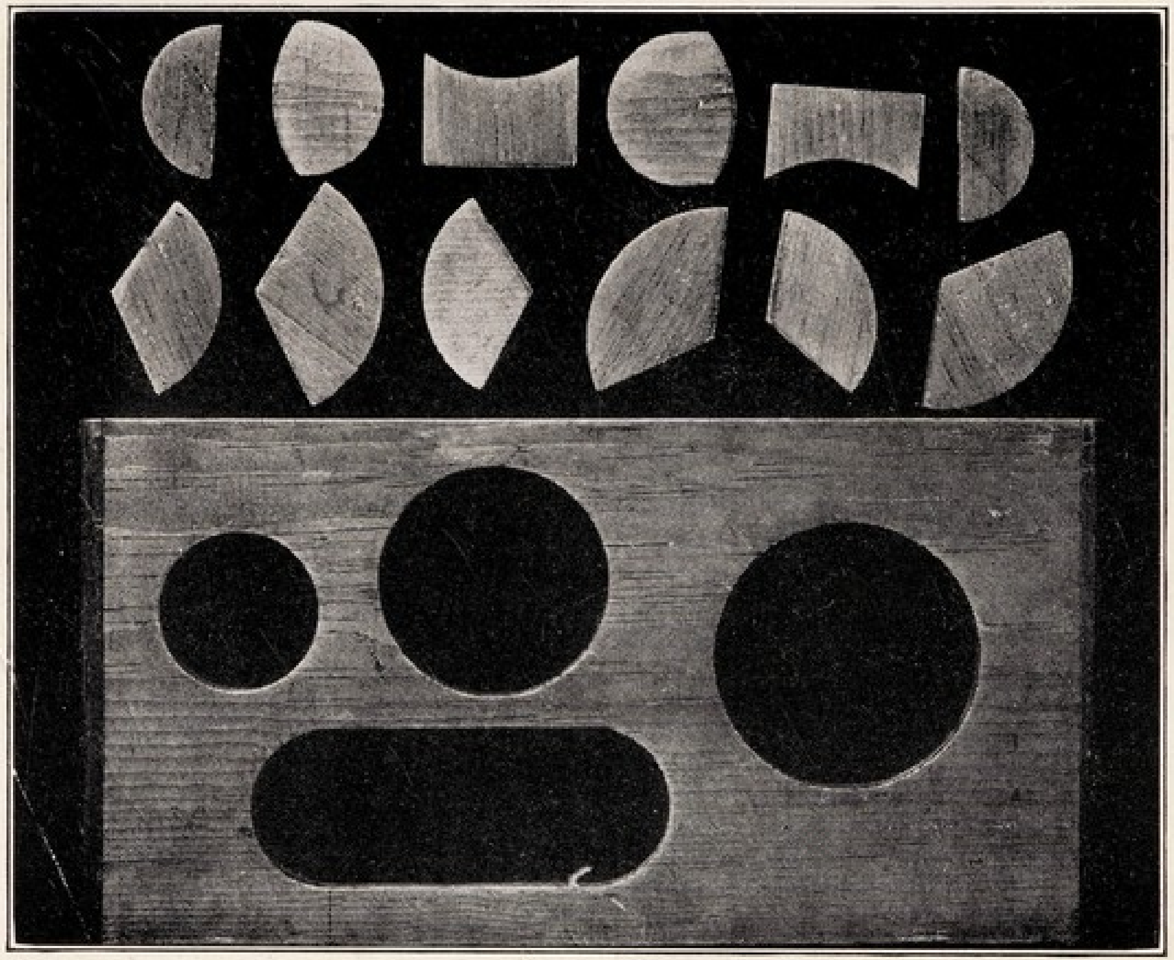


FIGURE 9. — The Casuist Form Board. (From Pintner and Paterson's *A Scale of Performance Tests.*)

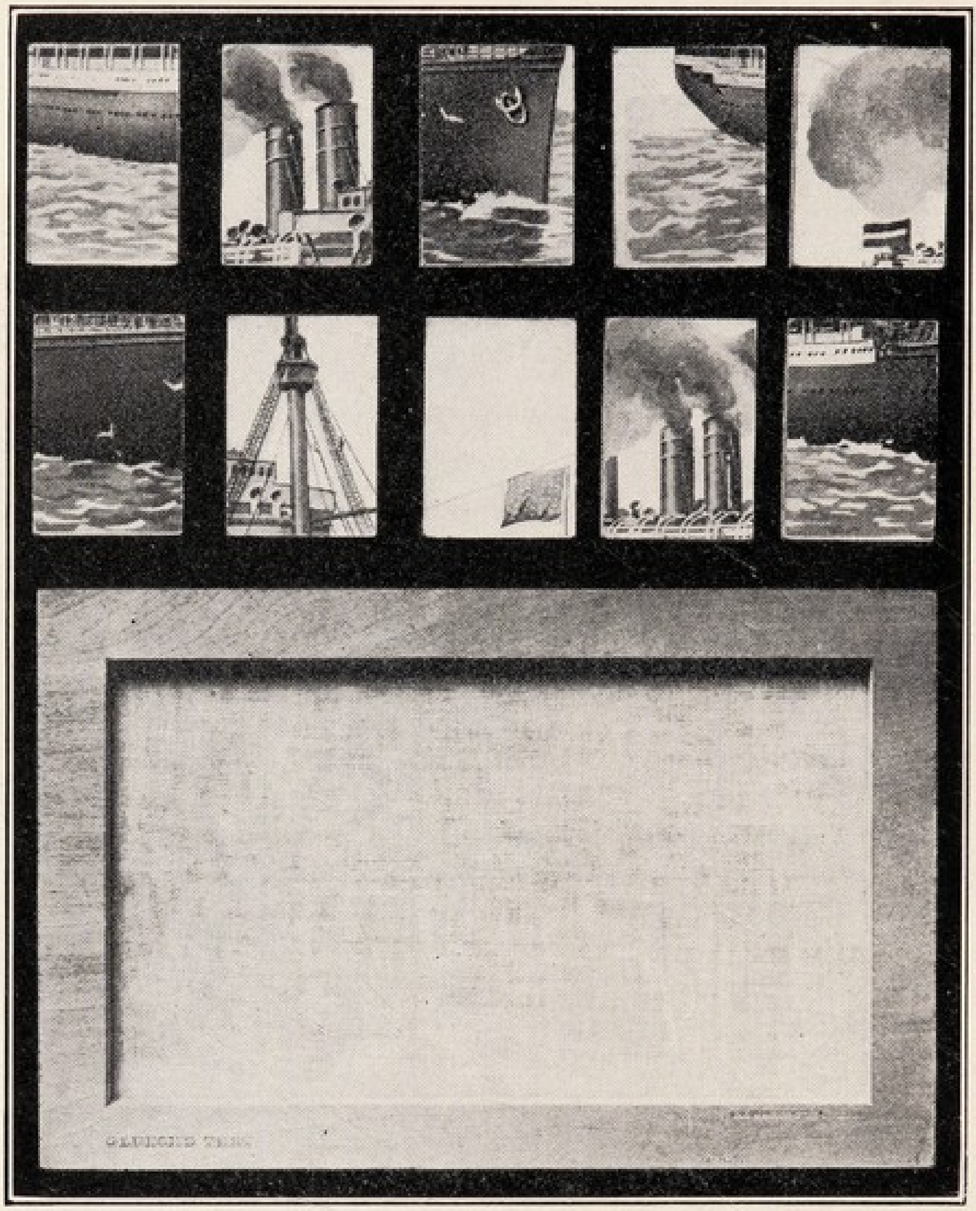


FIGURE 10. — The Ship Test. (From Pintner and Paterson's *A Scale of Performance Tests.*)

Concrete Intelligence. — The Performance Scale is measuring intelligence as shown in reactions to concrete situations, as contrasted with the more abstract type of situation presented in the Binet. The Performance Test requires no language and hence can be used with subjects who cannot understand English. The use of the Performance Scale along with the Binet enlarges our view of the intelligence of the subject. This scale gives the examiner a good opportunity to study the child at work without the constant interruption necessary with the Binet. The child's attitude and methods of work will frequently give the examiner valuable information about the non-intellectual traits of the child.

Difference between Individual and Group Tests. — The individual test method of examining a child requires special training on the part of the examiner. Each test must be given in exactly the same manner to each child, and yet each child must be treated differently in order to make sure that he is trying his best and giving his fullest cooperation. In most individual tests there are no rigid time limits, as we have in group tests. This flexibility of the circumstances surrounding the individual test, combined with the rigidity of the presentation of the actual test itself, makes the individual test strikingly different from the group test. An experienced psychological examiner will follow any useful leads which the reactions to the test situations on the part of the subject may reveal. He will do this, however, without in the least spoiling the standard presentation of the test itself. At the end of the examination, therefore, he will not only have a mental age determined in absolute conformity to the requirements of the scale itself, but he will also have obtained much valuable information about the subject which the mental age itself cannot reveal. This wider opportunity of observation that the individual test affords, makes it of paramount importance in the psychological clinic. A mere routine examiner will gain from his interview with the child nothing but a bare mental age, but an experienced clinical psychologist will have in addition, many valuable clues as to the various abilities of the child and as to his character and personality.

Group Tests. — In group tests this close contact between the examiner and the subject is not possible. The very nature of the examination prevents the possibility of making individual adaptations to the particular idiosyncrasies of the subjects. All the subjects, because they are in a group, are treated exactly alike. Directions for performing any particular test are given only once and are not repeated for the benefit of any particular subject, as may be done in many individual tests. If any subject is unwilling or shy or frightened or nervous or in a bad mood, he cannot be given special attention by the examiner and the test halted until better *rapport* has been established. Indeed the examiner is almost always unaware of such individual differences in the group that confronts him. Again the time limit for any test must necessarily be rigidly adhered to. All of these factors make the group test less flexible than the individual test, so that the mental rating resulting from a group test may not always be as good a measure of what a subject can do under the most favorable circumstances, as is the case with a well-administered individual test. On the other hand, the group examination spurs on many subjects to maximum effort because of the feelings of rivalry and competition which arise owing to the presence of others working at the same task. And, furthermore, just because it makes no allowance for individual feelings of laziness, shyness and the like, the mental rating resulting from a group test contains within itself a measure of the general adaptability and conformability of the individual to the school situation. In one sense, therefore, the group mental age is not as pure and undiluted a measure of intelligence as the individual mental rating is. Mixed in with the group intelligence rating is a rating of the docility of the individual, his willingness to follow directions, to cooperate in school work and work at his maximum under orders. This is no drawback to the group test for school purposes, because these non-intellectual factors enter into most school activities. The group test is, therefore, just as valuable as the individual test in helping us obtain a complete picture of the general intelligence of the individual.

There are all types of group intelligence tests suited for all levels of intelligence from the kindergarten stage to the university. The vast majority are verbal, either in the sense that they require a knowledge of understanding the spoken language of the directions, or else assume the ability to read printed words, or both. There are a few that are strictly non-verbal, not depending on language either in understanding directions or in working the test. Brief descriptions of some of these different types of group tests will be given. Only one or two tests of each type will be mentioned, the choice being made among commonly used tests. There are many other tests just as good as those selected as samples.

A Kindergarten or Primary Test. — The Pintner-Cunningham Primary Mental Test can be given to groups of children entering school for the first time, either in the kindergarten or first grade. It consists of a booklet of pictures and the children are told what to do with each set of pictures. Figure 11 shows page 4 of the test booklet. The children are told to "Look at the three girls at the top of the page. I want you to find the prettiest. Don't tell any one, but look at all the girls. Then put a mark on the prettiest girl. Go ahead. (Wait until all have finished). Now look at all the elephants. Mark the prettiest elephant. (Pause not longer than 10 seconds). Look at all the houses. Mark the prettiest house.

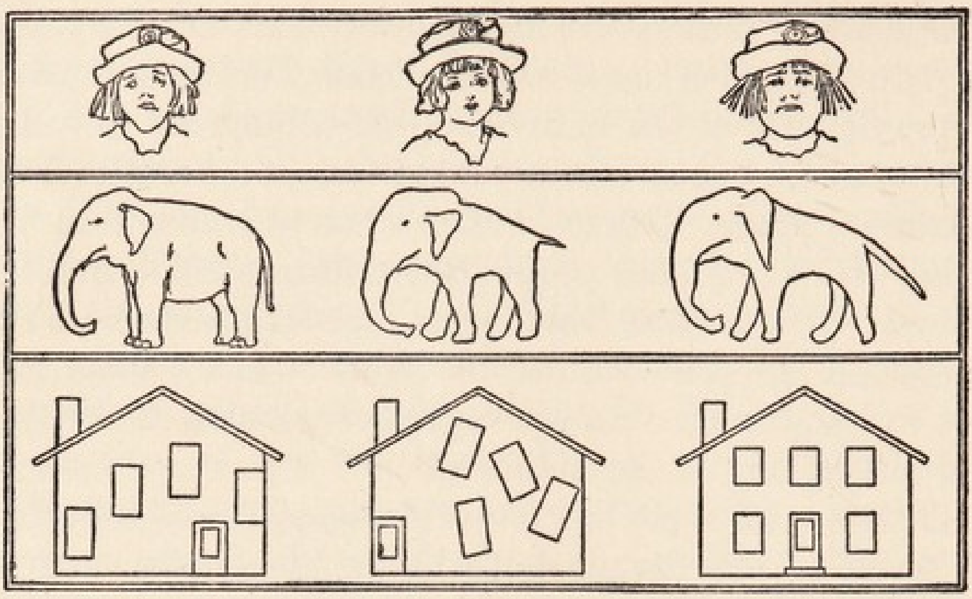


FIGURE 11. — Mark the prettiest in each row. From the Pintner-Cunningham Primary Mental Test. (World Book Co.)

(Pause 10 seconds.)" Figure 12 shows page 12 of the test booklet, and here the problem is to find the head for the man and mark it, and then the arm and then the leg.

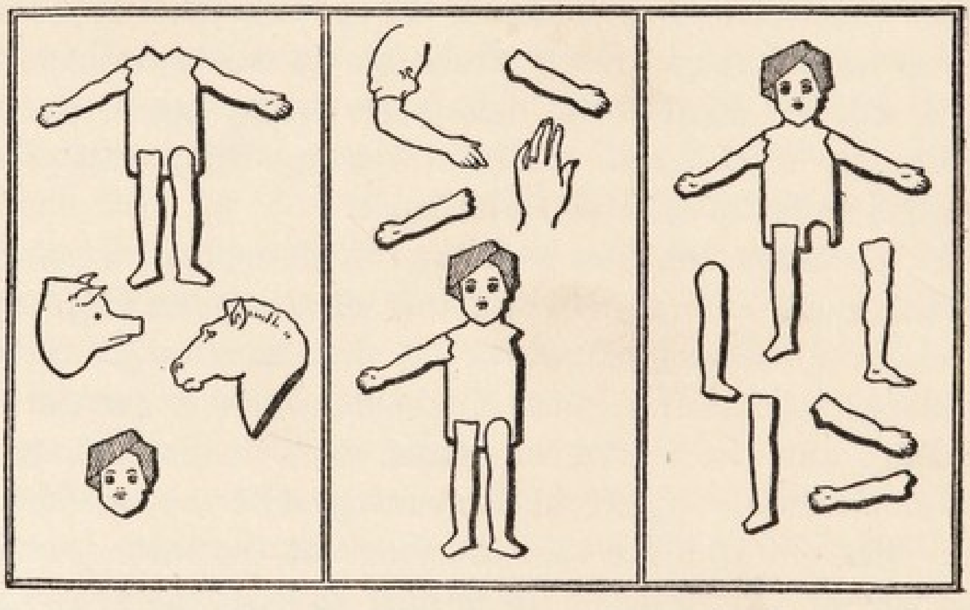


FIGURE 12. — Find the missing part. From the Pintner–Cunningham Primary Mental Test. (World Book Co.)

There are altogether seven sub-tests in the total test, with many items in each sub-test. None of the items require a knowledge of words or letters or numbers, but the children must be able to understand English in order to follow the directions. The test is given in the spirit of a game and young children are very much interested in it. A group test at this period of school life is extremely useful in helping the teacher to classify a new group of children about whose abilities practically nothing is known. In one sense this is the most crucial period of a child's school life, because his first year in school may determine his attitude towards school in general, and during the first year habits of industry or idleness are readily acquired.

An Elementary School Test. — The National Intelligence Test can be used with children from about the fourth to the eighth grade. Each test consists of five sub-tests, and the sub-tests for Scale A are: (1) Arithmetic Reasoning. (2) Completing Sentences, such as, "The dog —— black." (3) Common Attributes, e.g. underline the two words which tell what a mouse always has — back, cat, eyes, cheese, trap. (4) Similarities and Differences, e.g. marking which of pairs of words have similar or different meanings — new and old, fall and drop, expect and anticipate. (5) Symbol-Digit Test, in which the subject copies the digits for the appropriate symbols as shown in a key at the top of the test page.

Figure 13 shows Test 3 of the National Intelligence Test, Scale A.

The National Intelligence Test is very widely used in schools and has proved an extremely useful test. There are two scales and several forms of each scale.

A High School Test. — The Terman Group Test of Mental Ability is one of the most frequently used tests for high school pupils. It consists of ten tests as follows: (1) Information; underline the right word in "Coffee is a kind of bark, berry, leaf, root." (2) Best Answer; check the best answer to this question — Spokes of a wheel are often made of hickory because (a) hickory is tough, (b)

In each row draw a line under each of the two words that tell what the thing always has.

SAMPLES
- man (body cane head shoes teeth)
- dog (blanket chain collar legs nose)
- house (cellar paint room servants walls)

Begin here 1 hen (chickens corn feathers neck roost)
2 tiger (bones cage cubs fur jungle)
3 squirrel (acorn fur nest peanut tail)
4 kitten (ball claws eating eyes mouse)

5 bicycle (basket bell brake frame wheels)
6 stone (field hardness hurt throwing weight)
7 lion (cage head keeper mane prey)
8 face (cheeks eyebrow glasses mustache mouth)

9 forest (cones flowers grass soil trees)
10 paper (edges envelope printing surface watermark)
11 Borneo (airplanes land rivers saloons universities)
12 gully (flowers sand sides steepness stream)

13 piano (keys music pedals scarf stool)
14 satisfaction (conquering contentment money pleasure trouble)
15 illness (ailment discomfort doctor nurse recovery)
16 mob (confusion excitement hanging negro torches)

17 fire (ashes danger flame heat wood)
18 sea (coast reefs salt shoals submarines)
19 alley (cans fence narrowness passage pavement)
20 crime (death lawlessness punishment theft wrong)

21 pilot (cap chart knowledge license raincoat)
22 measles (discomfort doctor nurse rash recovery)
23 nun (beauty convent teacher vow woman)
24 citizen (city country male privileges vote)

FIGURE 13. — Common Attributes. The National Intelligence Test. Scale A. (World Book Co.)

it cuts easily, (c) it takes paint nicely. (3) Word Meaning; mark whether pairs of words mean the same or opposite, e.g. expel — retain; comfort — console; etc. (4) Logical Selection; underline the two right words in "A horse always has harness, hoofs, shoes, stable, tail." (5) Arithmetic Reasoning Problems. (6) Sentence Meaning; answer "yes" or "no" to such questions as: Does a conscientious person ever make mistakes? Is an alloy

ANALOGIES

SAMPLES	Ear is to hear as eye is to table see hand play Hat is to head as shoe is to arm coat foot leg

Do them all like samples.

---

1 Coat is to wear as bread is to
eat starve water cook. 1
2 Week is to month as month is to
year hour minute century 2
3 Monday is to Tuesday as Friday is to
week Thursday day Saturday 3
4 Tell is to told as speak is to
sing spoke speaking sang 4
5 Lion is to animal as rose is to
smell leaf plant thorn 5

6 Cat is to tiger as dog is to
wolf bark bite snap 6
7 Success is to joy as failure is to
sadness luck fail work 7
8 Liberty is to freedom as bondage is to
negro slavery free suffer 8
9 Cry is to laugh as sadness is to
death joy coffin doctor 9
10 Tiger is to hair as trout is to
water fish scales swims 10

11 1 is to 3 as 9 is to
18 27 36 45 11
12 Lead is to heavy as cork is to
bottle weight light float 12
13 Poison is to death as food is to
eat bird life bad 13
14 4 is to 16 as 5 is to
7 45 35 25 14
15 Food is to hunger as water is to
drink clear thirst pure 15

16 b is to d as second is to
third later fourth last 16
17 City is to mayor as army is to
navy soldier general private 17
18 Here is to there as this is to
these those that then 18
19 Subject is to predicate as noun is to
pronoun adverb verb adjective 19
20 Corrupt is to depraved as sacred is to
Bible hallowed prayer Sunday 20

Right . . . .

FIGURE 14. — The Analogies Test from the Terman Group Test of Mental Ability. (World Book Co.)

a kind of musical instrument? etc. (7) Analogies; Coat is to Wear as Bread is to eat, starve, water, cook. (8) Mixed Sentences; mark true or false mixed up sentences, such as, true bought cannot friendship be; etc. (9) Classification; cross out the word that does not belong in the series, e.g. Frank, James, John, Sarah, William; or death, grief, picnic, poverty, sadness. (10) Number Series; continue the series with two numbers, e.g. 8 7 6 5 4 3 — —; or 16, 8, 4, 2, 1, ½, — —. Figure 14 shows Test 7 of this test.

A College Test. — The most thorough-going test devised for college students is the Thorndike Intelligence Examination. It consists of a battery of four tests, the first two containing much of the same type of material we have already described and the last two emphasizing strongly the ability to read and understand difficult passages.

A Non-Language Test. — All of the tests so far described presuppose an English-speaking environment for the child. They are testing intelligence by means of verbal knowledge. The Pintner Non-Language Test is a good example of a very usable test which makes no such presuppositions. It does not use oral language in the directions nor printed words in the test itself. It consists of six tests as follows: (1) Movement Imitation, i.e. reproducing the movements of a pointer after it has been moved from dot to dot in different ways on the blackboard. (2) Easy Learning; (3) Hard Learning; these two tests are tests of the digit-symbol type involving the learning of new associations. (4) Drawing Completion, i.e. drawing in the missing parts of pictures. (5) Reversed Drawings, i.e. reproducing geometrical forms as they would be when turned upside-down. (6) Picture Reconstruction, i.e. indicating by numbers the positions of the parts of pictures so as to make a complete picture.

This test is well adapted to children in Grades III to

VIII. Figure 15 shows the drawing completion page of this test, where the problem is to complete the essential missing part of the picture.

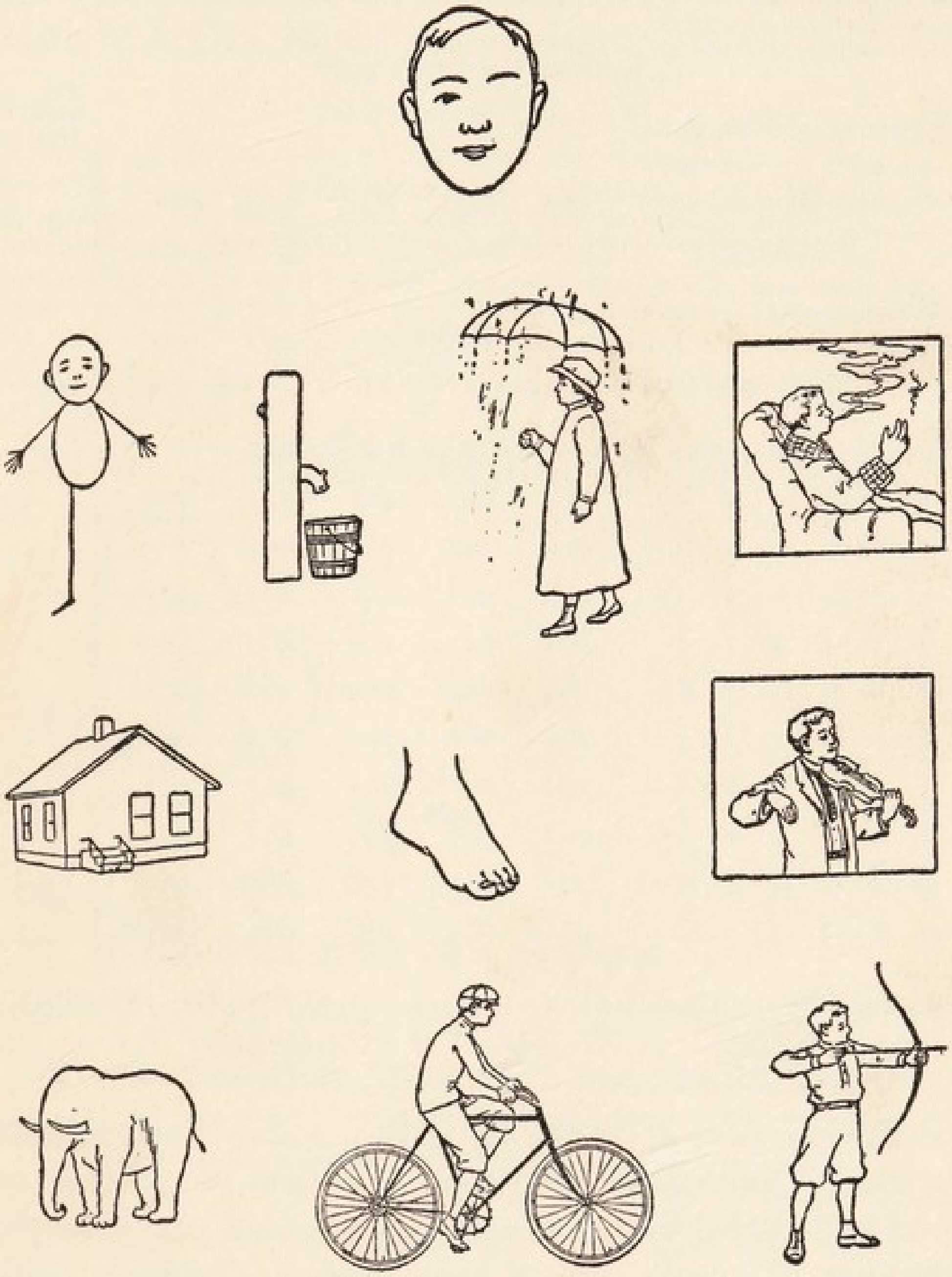


FIGURE 15. — The Picture Completion Test of the Pintner Non-Language Mental Test.

The great value of this non-language test is that it can be given to children of foreign parentage whose English is poor. It can also be effectively used to compare children of different nationalities, and it is one of the few group tests that can be used with deaf children. It is also very useful for all children, as it is testing a somewhat different aspect of intelligence from the usual verbal test.

NUMBER SEQUENCE TEST

Look at the sample that follows: — **Put letter here**

2 4 6 8 10 12—

What number should come next? (A) 17 (B) 13 (C) 14 (D) 15 (E) 9

The right answer is 14, so put the letter that goes with it, C, in the margin. ——→ C

Read this sample:

1 8 2 8 3 8— (E) 8 (F) 6 (G) 2 (H) 4 (J) 9

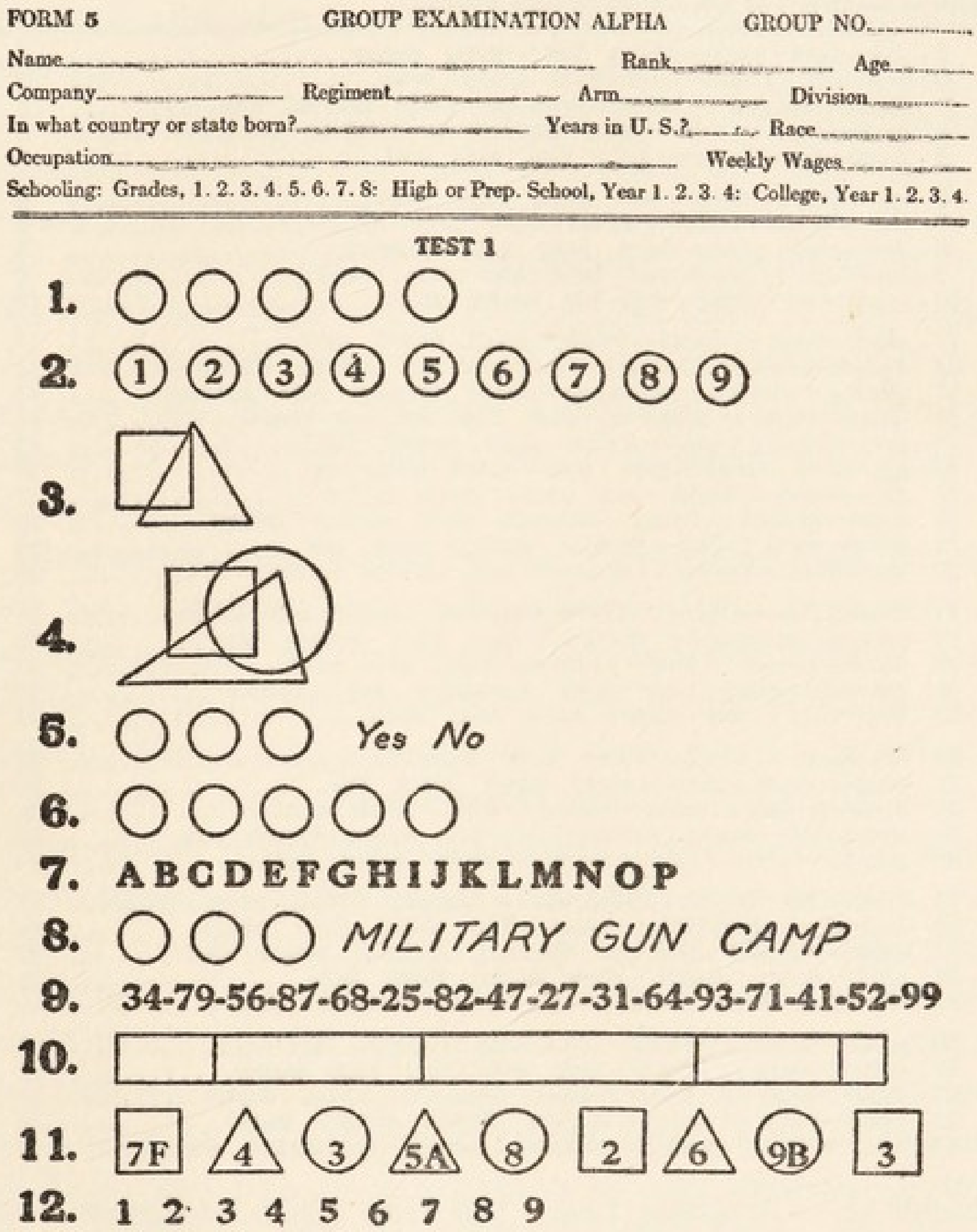
The right answer is 4, so you must put H in the margin. ——→ H

Do the rest in the same way. Remember to put the letter in the margin.

1	2	3	4	5	6—	(N) 5	(O) 2	(P) 7	(Q) 3	(R) 8	....
40	35	30	25	20	15—	(S) 10	(T) 5	(U) 14	(V) 15	(W) 20	....
8	7	6	5	4	3—	(W) 1	(X) 3	(O) 9	(Y) 2	(Z) 4	....
3	3	5	5	7	7—	(A) 8	(B) 11	(C) 9	(D) 6	(E) 10	....
5	9	13	17	21	25—	(F) 28	(G) 30	(H) 29	(J) 21	(K) 26	....
9	9	7	7	5	5—	(K) 4	(L) 10	(M) 6	(N) 2	(O) 3	....
4	3	5	4	6	5—	(K) 3	(L) 7	(M) 6	(N) 2	(O) 3	....
19	16	14	11	9	6—	(K) 3	(L) 5	(M) 2	(N) 7	(O) 4	....
103	95	87	79	71	63—	(D) 64	(E) 54	(F) 47	(G) 55	(H) 51	....
81	27	9	3	1	1/3—	(X) 2/3	(Y) 1/9	(Z) 0	(A) 1/27	(B) 1/81	....

FIGURE 16. — Test 3 of the Pintner Rapid Survey Intelligence Test. (T. C. Bureau of Publications.)

A Survey Test. — A Survey Test is a short test designed for survey purposes where large numbers of children are to be tested and where the emphasis is laid upon the performance of groups rather than upon individuals. The Pintner Rapid Survey Intelligence Test is a short test made up of four sub-tests of the usual verbal type, namely, (1) Opposites; (2) Analogies; (3) Number Sequence; (4) Classification. Figure 16 shows the Number Sequence Test in Form A of this test. Very rapid and accurate scoring is achieved on this test by the device of having the child write a number or letter in the margin, the correct numbers or letters forming a sequence easily remembered by the scorer. For example the right answers in the test shown in Figure 16 will spell the word PSYCHOLOGY.



FORM 5 GROUP EXAMINATION ALPHA GROUP NO.............

Name.......... Rank.......... Age..........

Company.......... Regiment.......... Arm.......... Division..........

In what country or state born?.......... Years in U. S.?.......... Race..........

Occupation.......... Weekly Wages..........

Schooling: Grades, 1. 2. 3. 4. 5. 6. 7. 8: High or Prep. School, Year 1. 2. 3. 4: College, Year 1. 2. 3. 4.

TEST 1

1\.

2\. 1 2 3 4 5 6 7 8 9

3\.

4\.

5\. Yes No

6\.

7\. A B C D E F G H I J K L M N O P

8\. MILITARY GUN CAMP

9\. 34-79-56-87-68-25-82-47-27-31-64-93-71-41-52-99

10\.

11\. 7F 4 3 5A 8 2 6 9B 3

12\. 1 2 3 4 5 6 7 8 9

FIGURE 17. — Directions Test. — The Army Alpha Intelligence Test.

Tests for Adults. — There are several group intelligence tests suitable for adults. The Otis Self-Administering Intelligence Test, Higher Examination, is one of the most

TEST 7

SAMPLES
- sky—blue : : grass— **table green warm big**
- fish—swims : : man— **paper time walks girl**
- day—night : : white— **red black clear pure**

In each of the lines below, the first two words are related to each other in some way. What you are to do in each line is to see what the relation is between the first two words, and underline the word in heavy type that is related in the same way to the third word. Begin with No. 1 and mark as many sets as you can before time is called.

1	shoe—foot : : hat— **kitten head knife penny**	1
2	pup—dog : : lamb— **red door sheep book**	2
3	spring—summer : : autumn— **winter warm harvest rise**	3
4	devil—angel : : bad— **mean disobedient defamed good**	4
5	finger—hand : : toe— **body foot skin nail**	5
6	legs—frog : : wings— **eat swim bird nest**	6
7	chew—teeth : : smell— **sweet stink odor nose**	7
8	lion—roar : : dog— **drive pony bark harness**	8
9	cat—tiger : : dog— **wolf bark bite snap**	9
10	good—bad : : long— **tall big snake short**	10
11	giant—large : : dwarf— **jungle small beard ugly**	11
12	winter—season : : January— **February day month Christmas**	12
13	skating—winter : : swimming— **diving floating hole summer**	13
14	blonde—light : : brunette— **dark hair brilliant blonde**	14
15	love—friend : : hate— **malice saint enemy dislike**	15
16	egg—bird : : seed— **grow plant crack germinate**	16
17	dig—trench : : build— **run house spade bullet**	17
18	agree—quarrel : : friend— **comrade need mother enemy**	18
19	palace—king : : hut— **peasant cottage farm city**	19
20	cloud-burst—shower : : cyclone— **bath breeze destroy West**	20
21	Washington—Adams : : first— **president second last Bryan**	21
22	parents—command : : children— **men shall women obey**	22
23	diamond—rare : : iron— **common silver ore steel**	23
24	yes—affirmative : : no— **think knowledge yes negative**	24
25	hour—day : : day— **night week hour noon**	25
26	eye—head : : window— **key floor room door**	26
27	clothes—man : : hair— **horse comb beard hat**	27
28	draw—picture : : make— **destroy table break hard**	28
29	automobile—wagon : : motorcycle— **ride speed bicycle car**	29
30	granary—wheat : : library— **read books paper chairs**	30
31	Caucasian—English : : Mongolian— **Chinese Indian negro yellow**	31
32	Indiana—United States : : part— **hair China Ohio whole**	32
33	esteem—despise : : friends— **Quakers enemies lovers men**	33
34	abide—stay : : depart— **come hence leave late**	34
35	abundant—scarce : : cheap— **buy costly bargain nasty**	35
36	whale—large : : thunder— **loud rain lightning kill**	36
37	reward—hero : : punish— **God everlasting pain traitor**	37
38	music—soothing : : noise— **hear distracting sound report**	38
39	book—writer : : statue— **sculptor liberty picture state**	39
40	wound—pain : : health— **sickness disease exhilaration doctor**	40

FIGURE 18. — Analogies Test. The Army Alpha Intelligence Test.

useful of these. The Army Alpha Test is another very useful test. It was one of the first group tests devised and was used extensively during the World War in testing about two million men. As one of the first group tests to be devised, it has influenced many group tests that have since been constructed. Almost all of the verbal tests at the present time show distinct resemblances to the original Army Alpha Test. Figures 17 and 18 show test 1 and test 7 of the Army Alpha. Test 1 shown in Figure 17 is a test of Following Oral Directions. The first direction is "Make a cross in the second circle, and also a figure 1 in the third circle." The fourth direction is, "Make a figure 1 in the space which is in the triangle but not in the circle or square, and also make a figure 2 in the space which is in the square and circle, but not in the triangle." The seventh direction is "Cross out the letter just before D and also draw a line under the second letter before I." The last direction is, "If 6 is more than 4, then (when I say 'go'), cross out the number 5 unless 5 is more than 7, in which case draw a line under the number 6."

Evaluation of Scores. — We have described briefly one test for each level of intelligence from kindergarten to college, one non-language test and one survey test. There are many other tests and we do not mean to indicate by the selection made that these are the best tests. There are many others equally good. After any one of these group tests has been given, there results a score for each person examined. This score may be then converted into some type of mental rating which interprets the score. The most common method is to convert the score into a mental age by means of norms or standards for that purpose. When a mental age has been obtained, an I.Q. may be calculated in the usual way. This use of M.A. and I.Q. is a following of the method devised in conjunction with the Binet-Simon Tests. It is undoubtedly useful and easy to understand, and is, moreover, fairly well understood by teachers in general. There is, however, a danger in presupposing that the M.A. on all tests means exactly the same thing, and much surprise and astonishment may occur when the same child is found to have different mental ages on different tests. As we shall see later this is perfectly possible, if the different tests are measuring different aspects of intelligence. Furthermore, the difference in the control over the individual exercised by the examiner in the individual and group tests respectively may lead to differences in the results. In general one feels that the M.A. resulting from a thorough individual test is more indicative of the all-round ability of the child than the M.A. resulting from one group test. With high school and college students it is difficult, if not impossible, in most cases to obtain a reliable M.A. by means of a group test. Mental ages above 14 or 15 become fictitious. In these cases, therefore, other types of mental rating are resorted to. Probably the best method for all kinds of tests for children of all ages is to express the score of the child in terms of the standard deviation of some known group. In this way we get the T score proposed by McCall and the Mental Index used by Pintner. All these kinds of ratings show us how much above or below the average a particular child stands. One may also convert the score on a test into a letter rating such as A, B, C and so forth. This was done in the Army Intelligence Testing and is used in several group tests. Each letter rating is defined as to its significance, usually in terms of the percentage making similar scores. Each of these different ways of expressing a mental rating possesses its special advantages for particular purposes.

Tests for Very Young Children. — None of the tests so far described are suitable for very young children. The group tests do not go below the kindergarten level and the Stanford Revision of the Binet Tests stops at age three.

The Kuhlmann Revision of the Binet (2) is valuable because it contains tests for three months, six months, twelve months, eighteen months and two years.

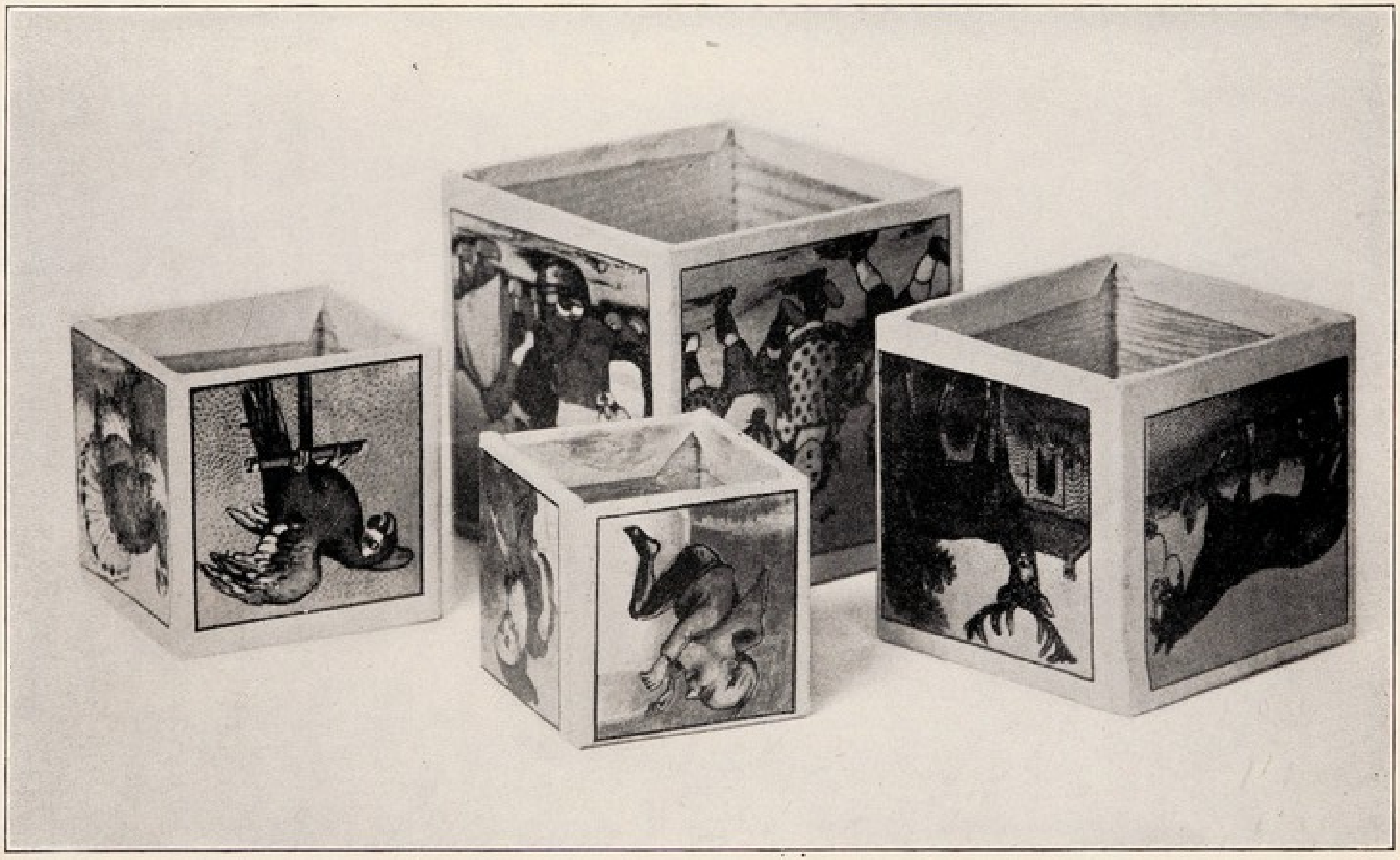
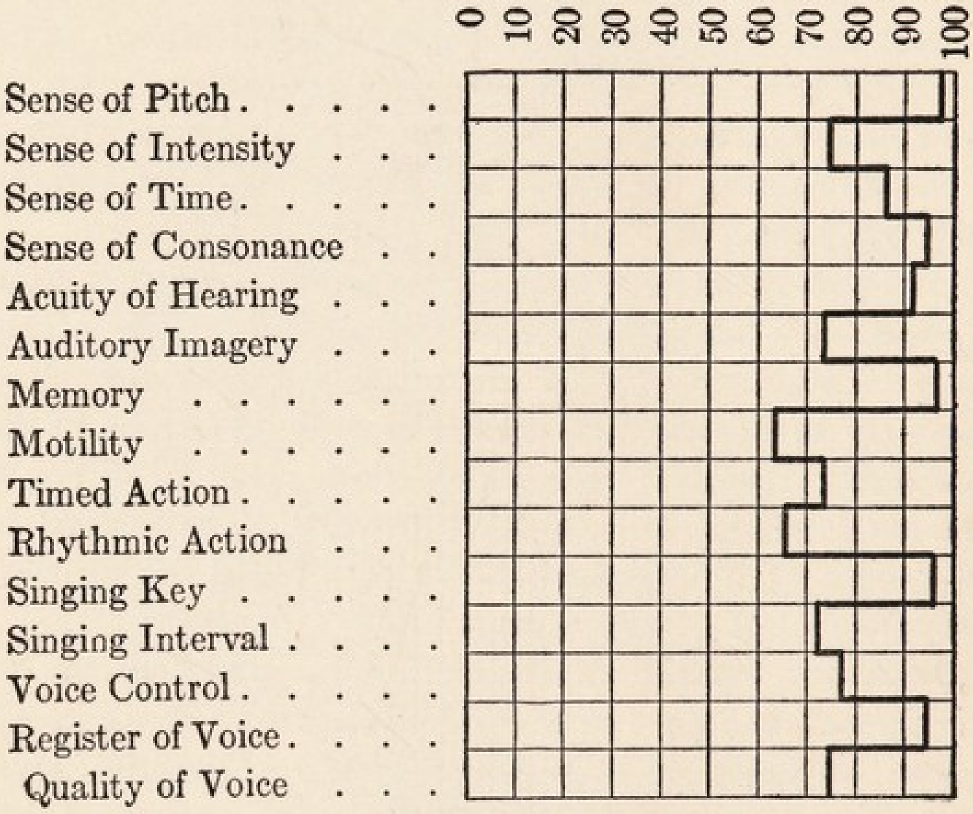


FIGURE 19.—The nest of cubes test from the Merrill Palmer Tests. (Courtesy of Rachel Stutsman of the Merrill Palmer School.)

The Merrill-Palmer Tests (6) consist of a great number of tests, mostly of the performance type, which can be given to children from 18 to 72 months. Some of the tests are: (1) Commands. "See the box. Take the box from the table and put it on the chair." (2) Straight Tower. Build a straight tower, three blocks high, making sure child watches the process; then say, "Now, you make one." (3) Questions. "What does a doggie say?" "What does a kittie say?" "What is this?" (pencil). "What is it for?" and so on. (4) Nest of cubes. Figure 19 shows the cubes taken apart. The nest is shown the child, it is taken apart, then put together again while he is watching, and then he is allowed to try. In addition there are form boards and picture puzzles.

Like the Binet Tests, we have a miscellaneous group of tests, with the performance type predominating. They are not non-language in character, however, as all of the tests are dependent upon the oral directions of the examiner. The tests are arranged in order of difficulty and are grouped in six-month intervals. The mental development of the young child is very rapid during these early years.

Tests for Special Capacities. — In addition to the attempt on the part of the psychologist to measure the general capacity or intelligence of the individual, there exist also attempts to measure special capacities. If special capacities exist and are to some extent at least inherited, it would seem possible that we might measure their strength in the individual. As to the number of such special capacities or talents, there is much doubt. We do not know whether such things as mathematical ability, literary ability, artistic ability and the like, are due to the interests and opportunities of the individual or whether they are dependent upon the inheritance of certain specific abilities in addition to the general ability or intelligence With regard to musical aptitude, however, there are good reasons for believing that it is dependent upon the inheritance of a specific capacity, and tests have been devised by Seashore to measure this capacity apart from such modifications as may have been produced by musical training. Figure 20 shows the results of these musical tests given to a highly gifted individual, and from this chart we can also see the several different characteristics that have been measured.



Talent Chart of Mr. White.

FIGURE 20. — The Measurement of Musical Talent. (From Seashore's *Psychology of Musical Talent*, Silver, Burdett and Company.)

Tests of mechanical ability have also been constructed, both of the pencil and paper type and also of the performance type, and these give us a measure of an individual's knowledge of and ability to handle and understand mechanical things.

The Distribution of Intelligence. — Whenever intelligence tests are given to a large unselected group of indi-

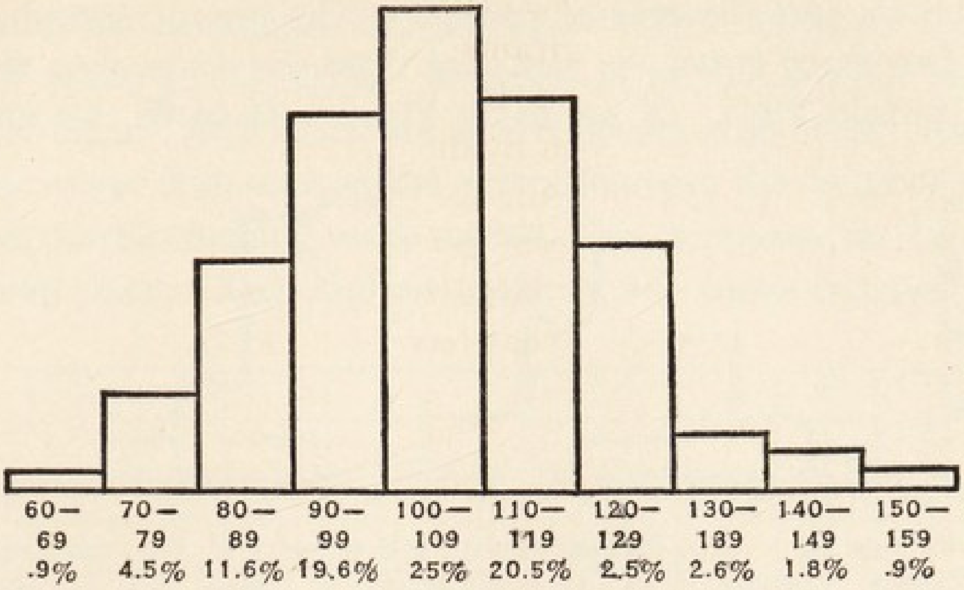


FIGURE 21. — I.Q. distribution of 112 Kindergarten Children. (From Terman's *The Intelligence of School Children.*)

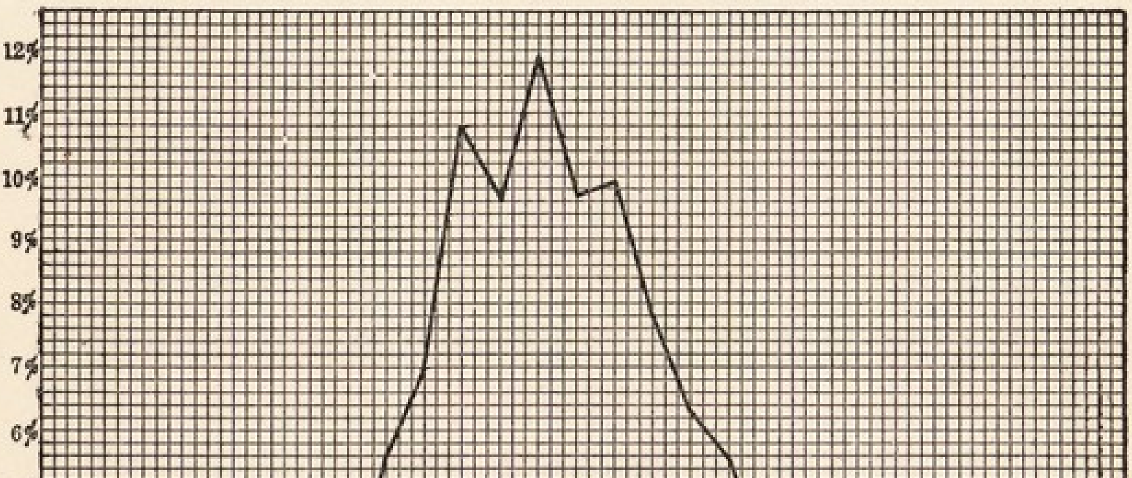


FIGURE 22. — Showing distribution of National Intelligence Test I.Q.'s of 1225 pupils above Grade IV, Public School 64. (From Irwin and Marks' *Fitting the School to the Child.*)

viduals and the mental ratings of the group plotted on a frequency curve, we find that these curves tend to take a certain form. If we study Figures 21 to 28, we note

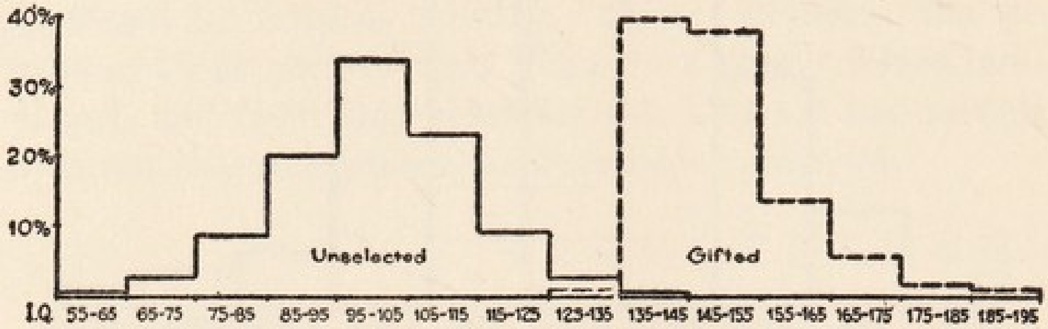


FIGURE 23. — I.Q. distribution for 999 gifted and 905 unselected children. (From Terman's *Genetic Studies of Genius,* Vol. I. Stanford University Press.)

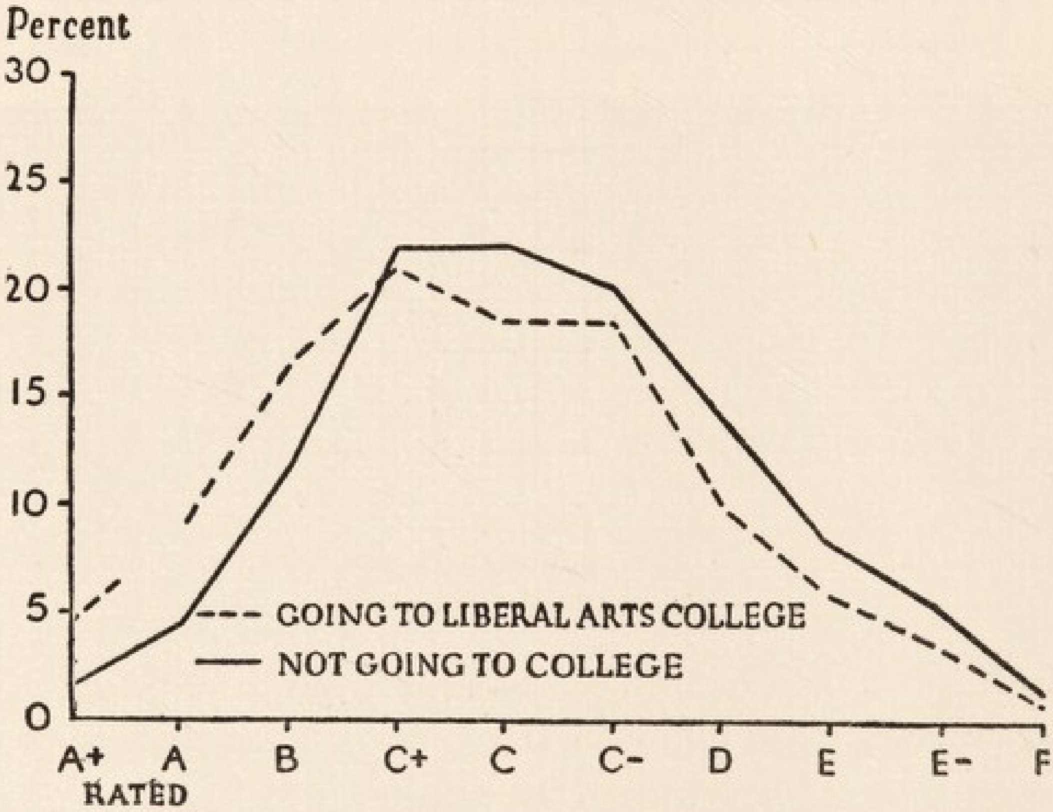


FIGURE 24. — Frequency curves for high school seniors. (From Book's *The Intelligence of High School Seniors.*)

that the general shape of the curve is more or less the same in all of them. The curves begin low at the left, rise gradually higher and higher until they reach their maximum in the middle and then gradually sink again to a low point at the right. In terms of an intelligence test this means that there are a few people who make very low scores and that as the scores increase the number of individuals making such scores also increases up to a given point. This highest point on the curve represents

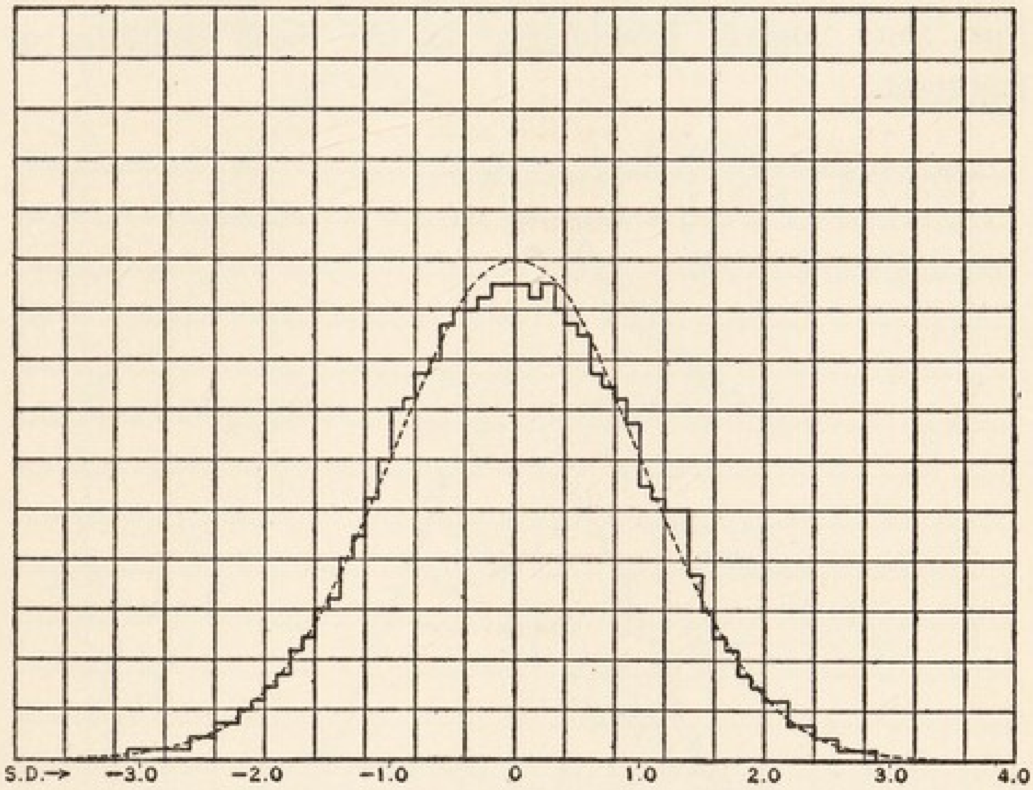


FIGURE 25. — Composite curve for the ninth grade based upon eleven tests. The broken line indicates the theoretical normal curve. (From Thorndike's *Measurement of Intelligence.*)

average or normal ability, because a great number of individuals cluster around this point. From this point on, as the scores increase, the number of individuals becomes less and less until we end up with few individuals making extremely high scores. In other words, the frequency of very low and very high intelligence scores is small. There are few idiots and few geniuses in the world. The frequency of the intermediate levels is somewhat greater. There are more dullards and bright people. Most numerous of all are the people of average intelligence, who cannot be called either dull or bright and who are far from being idiots or geniuses. Normal intelligence is that degree of ability possessed by the median or average individual, and, because intelligence is distributed as we have described, it follows, therefore, that normal intelligence is the most common or frequent.

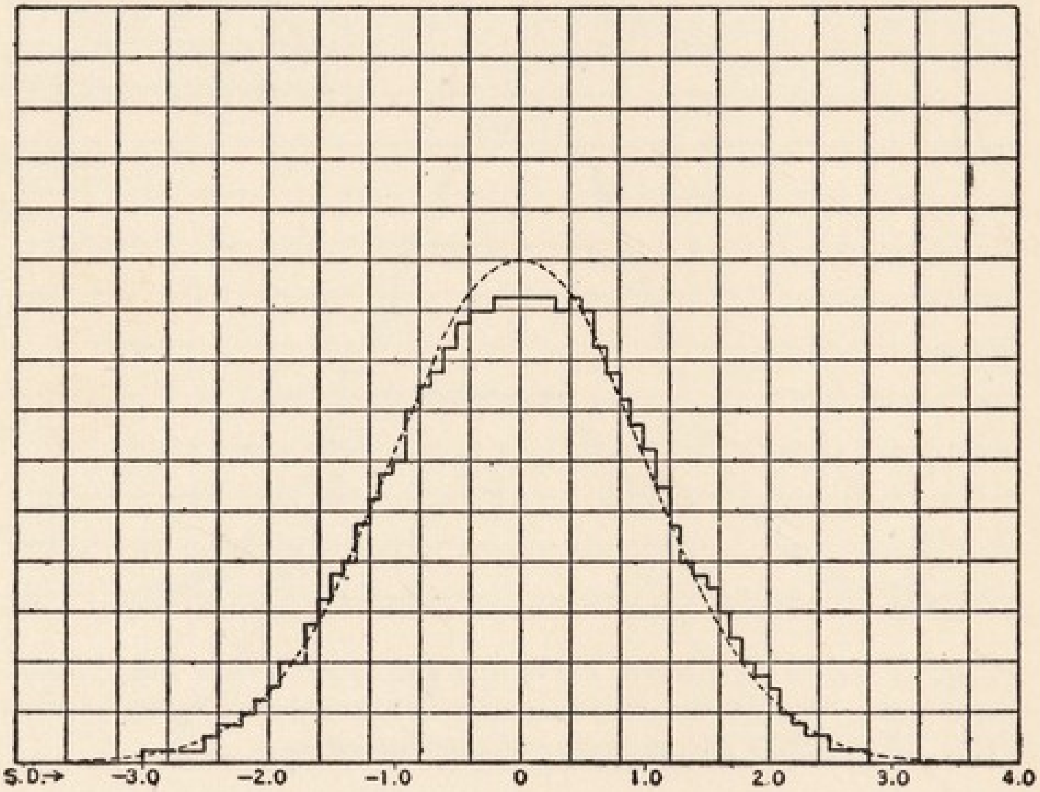


FIGURE 26. — Composite curve for college freshmen, derived from eleven single curves. The broken line indicates the theoretical normal curve. (From Thorndike's *Measurement of Intelligence.*)

The type of distribution we have been describing is generally called a normal distribution and the curves shown in Figures 21 to 28 are generally referred to as normal curves. Whenever a large number of individuals

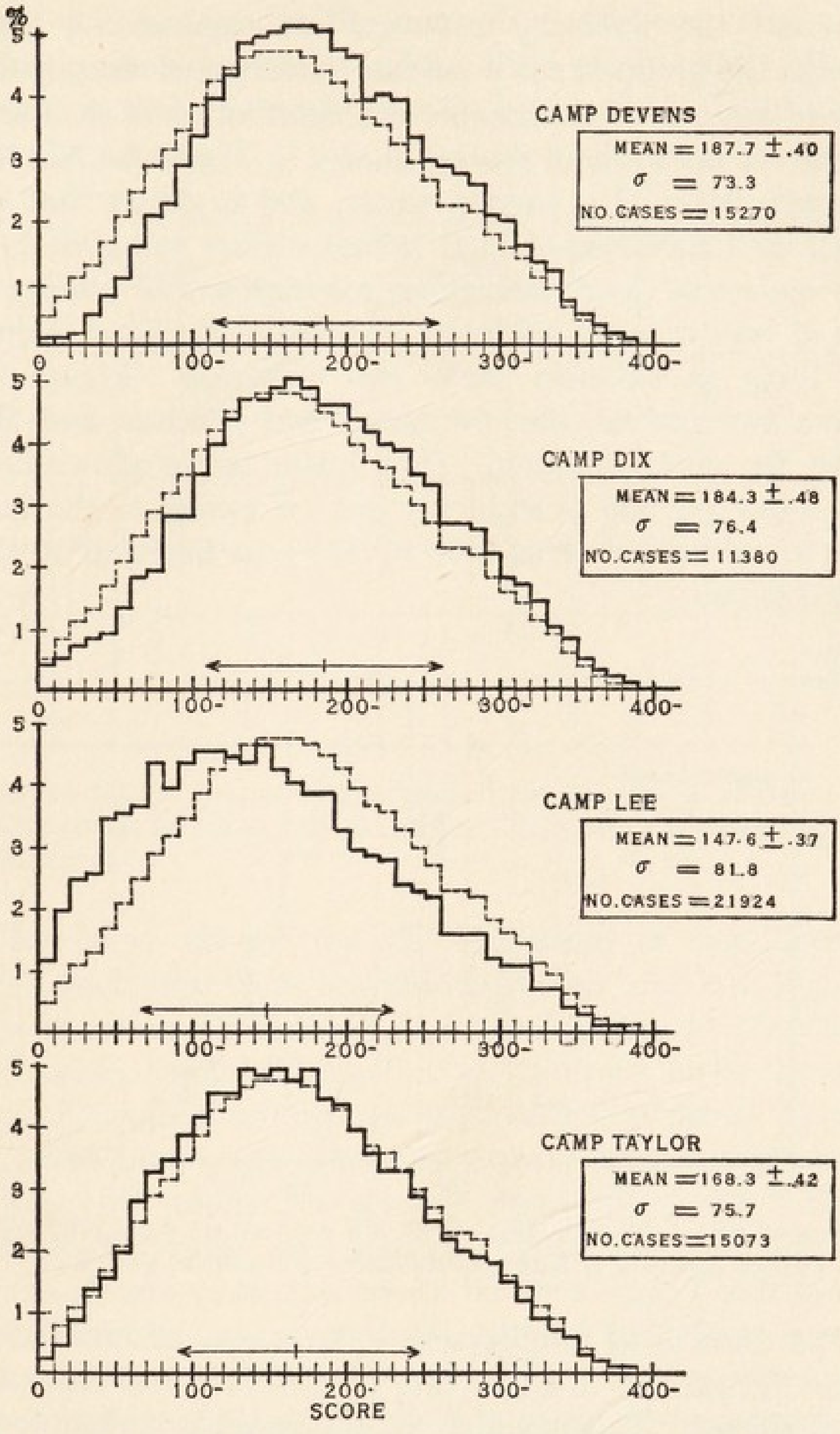


FIGURE 27. — Distribution of intelligence scores of white enlisted men. The dotted curve represents the average percentage distribution for the four camps. (From *Memoirs of the National Academy of Sciences,* Vol. XV.)

is tested the tendency is to get a normal curve, even though the group is not a random selection of the population in general. For example the distribution of the intelligence of high school seniors shown in Figure 24 has the characteristics of a normal curve, and it shows that an unselected sampling of high school seniors will give us a few seniors of poor intelligence for high school seniors, a few of very exceptional intelligence with the great majority lying in between these two extremes. Figure 23 shows two curves, one for unselected children and the other for gifted children. The latter curve shows how the lower portion is suddenly cut off owing to the fact that no child below an I.Q. of 135 was included in the gifted group.

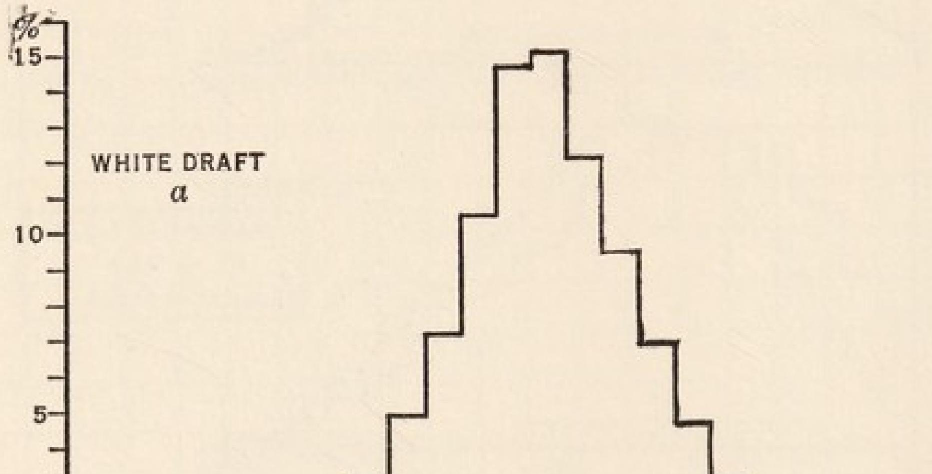


FIGURE 28. — Distribution of 93,965 soldiers on combined scale. (From *Memoirs of National Academy of Sciences*, Vol. XV.)

The Growth of Intelligence. — Everyone knows that as a child grows, he becomes able to do more and more difficult things. Things which were impossible for him to do or understand at three years, are readily done or grasped at eight. The capacity for reacting to more and more complex situations increases as the nervous system develops and grows. At the same time we know that the rate of growth diminishes as the child grows older. This general growth of intelligence is pictured by Thorndike (8) in Figure 29.

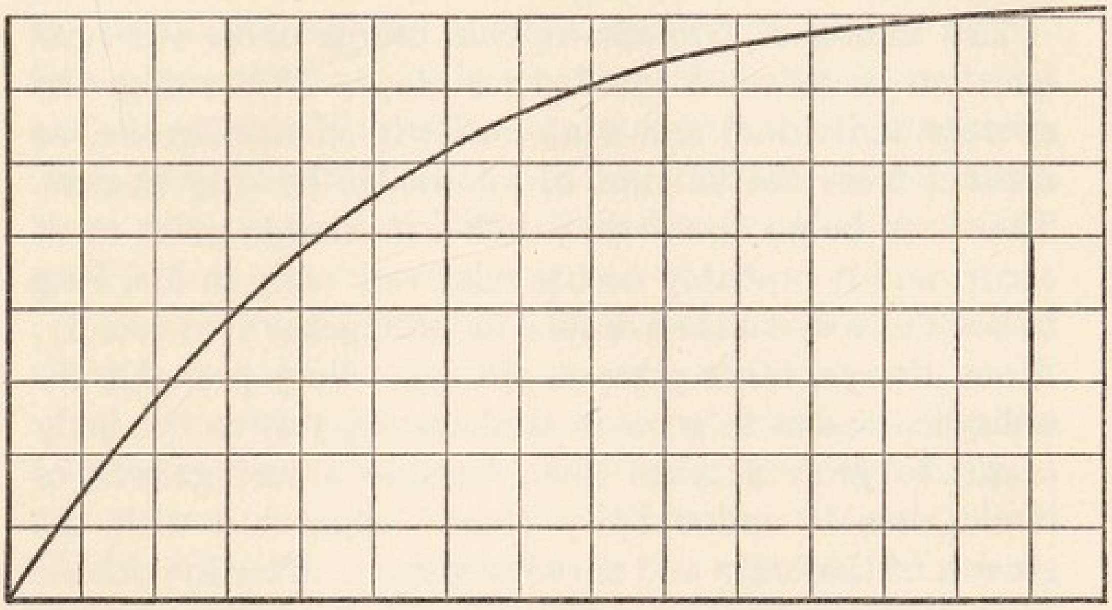


FIGURE 29. — The general nature of the relation of altitude of intellect to age in years, 0 to 20. (From Thorndike's *Measurement of Intelligence.*)

This curve shows that the increment of intelligence added each year becomes gradually less and less as the child increases in chronological age. Common observation would seem to bear this out. From birth to one year, the mental change, as well as the physical, in an infant is enormous. One can almost literally see him growing and learning to do new things every day. This marvellously rapid rate gradually slows up. We do not note such tremendous changes between ages 7 and 8, or ages 9 and 10. Some have maintained that there is another definite spurt at adolescence, but this has not been clearly demonstrated. Again it is easier to find intellectual tasks which will discriminate between the 5 and 6 year old than to find such as will discriminate between the 12 and 13 year old. Such general considerations as well as the results of actual intelligence testing lead us to believe that the curve of growth of intelligence is of the type pictured in Figure 29. The increment of intelligence added each year becomes less and less until it ceases altogether.

The Limits of Growth. — This brings us to the next question as to when intelligence stops. When does the average individual arrive at maturity of intelligence, as distinct from the amount of knowledge he may possess. There can be no doubt that such a maximum point must occur, and it probably occurs relatively early in life, long before the very marked decline in intelligence that usually, if not always, takes place in old age. Very probably intelligence ceases to grow in adolescence, just as the body ceases to grow at that time, because actual growth of intelligence is undoubtedly closely connected with the growth of the brain and nervous system. This hypothesis of the early cessation of the growth of intelligence does not mean, that thereafter an individual cannot learn anything new, just as a cessation of the growth of the body does not mean that thereafter no new physical accomplishments may be learned. A man may learn to skate or play golf or billiards later in life, and so also may he learn a foreign language or take up a new study. In each case, however, his success will be conditioned by the adequacy of his physical organization and the adequacy of his mental organization.

At present our intelligence tests do not measure much, if any, intelligence growth beyond the ages of 14 to 18 for the average individual. At just what age growth ceases is still debatable. Furthermore, we must remember that our intelligence tests are still rather coarse instruments of measurement, and that there may be growth after this period which is too slight to be detected by the instruments which we possess. In the army testing no difference in intelligence was found between the different age groups from twenty to thirty, so that intelligence evidently is not increasing between twenty and thirty. Furthermore average fourteen-year-olds score as much as the average army recruit. Tests given to unselected samplings of fourteen- and fifteen-year-olds by one investigator showed no increase in intelligence by the latter group. Other investigators think, however, that they have found a slight increase up to about age eighteen. A great difficulty arises because it is very difficult to obtain unselected samplings of children at ages above the elementary school.

The high school acts as a selective agency for children of higher intelligence. At present, therefore, we must leave as unsettled the actual point at which intelligence ceases to grow, the probability being that the point is somewhere between 14 and 18. We must remember that all this discussion has related to the average individual, for there is reason to believe that individuals possessing more than average intelligence, may continue growing in intelligence for a longer period, and, conversely, individuals below normal intelligence may cease growth earlier.

The Constancy of Growth. — Closely connected with the preceding question is the further one of the constancy or regularity of growth. Does the growth of intelligence of the average child continue steadily from year to year or are there marked fluctuations from time to time? There may be and probably are fluctuations, just as there are fluctuations in physical growth, but in general the picture would seem to be one of steady and constant growth. The evidence here is obtained from repeated tests of the same children. If the I.Q.'s from test to test remain relatively constant, then we may say that growth continues regularly. Here are the results for one child tested nine times by the Stanford Revision of the Binet:

Test	*C.A.*	*M.A.*	*I.Q.*
1	2–3	3–6	155
2	3–1	4–6	146
3	4–0	5–8	142
4	5–2	7–4	142
5	6–0	8–6	142
6	7–2	9–11	138
7	8–2	11–2	137
8	8–11	13–0	146
9	10–0	14–3	143

Probably the tests at the earlier ages on the Stanford are slightly too easy and hence the abnormally high I.Q. when the child was two years old. Apart from this one high I.Q., the others range from 137 to 146. This amount of variation over a five-year period is relatively small. In such cases, therefore, we say that the I.Q. is constant. In sixteen investigations of the I.Q.'s of varying numbers of children who had been given at least two tests, the median difference in I.Q. from the first test to the second is 4.6 points.

The amount of constancy or agreement from one test to another can also be expressed as a coefficient of correlation. The median correlation of seventeen such correlations found by different investigators is .90, and this means that in general there is a distinct tendency for the I.Q. of a child to remain fairly constant from one test to another. This is all that the psychologist means when he talks about the constancy of the I.Q. He does not mean that the I.Q. never shifts from year to year. He does not mean that there are not cases of very great difference from one test to the next. Cases showing changes of 20 or more points occur. These may be due to poor tests, poor testers or to actual changes that have taken place in the child. As to the causes of such real changes in the rate of growth of intelligence in some few children we are in ignorance. In general, however, the picture of constant steady growth with no violent fluctuations is the common one that has been found, and this is what the psychologist means when he says that the I.Q. is constant.

The Effect of Practice. — Familiarity with intelligence tests or practice or coaching leads, of course, to an increase in score. There is nothing magical about intelligence tests so as to make them impervious to the universal laws of learning — the laws of exercise and effect. If, therefore, we deliberately teach a child the items of an intelligence test, or if we give him the same intelligence test once every day for several successive days, we will obtain a score much higher than would have been obtained without such practice or coaching. This score will not be indicative of his intelligence, if it is interpreted in terms of M.A. or I.Q. as derived from norms established on uncoached children. We must remember here that our mental ratings are always comparisons with what the standard or normal group of children has done. If, therefore, a child deviates markedly from this group, such as our coached or practiced child, then we cannot legitimately interpret his score in terms of the standard group.

If children are given direct coaching on the Stanford Revision of the Binet, their mental ages can be increased about two years. This has been demonstrated by one investigator (Graves, 1). Why then does this not invalidate the Binet Scale as a measure of intelligence? Because the test is rarely given to children more than once a year. When so given the child does not always do the same tests, but as he grows in intelligence new and harder ones enter to test his ability. The relative stability of the I.Q.'s of children so tested shows that the practice gained by yearly testings is negligible. The example given above of the child given nine annual tests shows no effect of practice or else the I.Q. would have steadily increased. With a little care, therefore, the error that would be introduced by practice or familiarity or coaching may be easily avoided in the administration of intelligence tests.

The Description of Intelligence. — As we suggested at the beginning of this chapter, we have left any attempt to describe or define intelligence until the end. Our survey of the tests and the discussion of distribution and growth have probably helped the reader to grasp the psychologist's meaning of intelligence far better than any formal definition. Through and by means of our work with intelligence tests, we are slowly coming to a better and fuller meaning of what intelligence is. We may at present amplify our concept of intelligence by thinking of different kinds of intelligence and by considering different aspects of intelligence.

Kinds of Intelligence. — The use of different instruments of intelligence measurement has brought to our attention the fact that there are different kinds of intelligence, such as abstract, concrete and social. Abstract intelligence would signify the ability to respond to symbols of various sorts, such as words, numbers, letters and the like. Abstract intelligence is required by all who would progress in the ordinary academic subjects in school, such as reading, writing, arithmetic and so on. In its higher levels it is typified by the student and philosopher who is dealing with the relations of things symbolized in words or numbers.

Concrete intelligence denotes our ability to respond to concrete things. It is not mere skill in manipulation such as may come through practice, but it is rather the ability to understand concrete situations and react adequately to them. Performance tests measure this kind of ability and in everyday life such ability is shown in our reactions to doors, chairs, tables, lamps, stoves, street-cars, automobiles, steps, elevators, turnstiles, cash-registers, typewriters and the innumerable machines, instruments and things that people in different walks of life handle every day. In its higher levels it is typified by the surgeon, the inventor, the scientist, in so far as they are reacting to concrete things. No doubt these higher levels cannot be reached without also having a high, if not very high, level of abstract intelligence.

Social intelligence is probably a third kind of intelligence, although for its measurement we have as yet no separate scales. The assumption of this third kind of intelligence, is, therefore, hypothetical. We can, however, readily think of different degrees of ability in reacting to persons as distinct from concrete things and from abstract symbols. Ability to understand and react adequately to persons would be the criterion of social intelligence. This would not include the feelings or emotions aroused in us by other people, but merely our ability to understand others, to react in such a way toward them so that the ends desired should be attained. Some people seem frequently to react unfittingly to the presence of others and they may be thought of as possessing little of this kind of intelligence. Others on the contrary are very able in dealing with people. We say they are able to handle men well. Good generals, executives, teachers, politicians, preachers must all possess large amounts of this kind of intelligence.

This three-fold division of intelligence is merely a convenient scheme for dividing the numerous different reactions that may be taken as indicative of intelligence. There is no reason why other divisions might not be made, and, indeed, it is quite common to speak of verbal as opposed to non-verbal intelligence, and when we do this we are cutting across the previous scheme of classification and considering all reactions as divisible into verbal and non-verbal. Or rather, more narrowly, we are thinking of the difference between the ability to react to verbal test items as contrasted with the ability to react to non-verbal test items. Furthermore, in the above suggested three-fold classification of intelligence there is no suggestion of one kind of intelligence being contrary to or excluding the other. It is most probable that there is a high correlation between the three kinds. Great social intelligence presupposes a much above average abstract and concrete intelligence. The great executive or politician cannot be an imbecile or moron in abstract intelligence. So the great surgeon or inventor must have high abstract intelligence, and the great philosopher is much above average in his reactions to people and things.

Aspects of Intelligence. — From a different point of view we can extend our concept of intelligence by describing three different aspects which it possesses, namely speed, level and range. The conception of speed is simple and easily measurable on our tests by taking the time or by allotting a given amount of time for a certain number of tasks. Level is the height or altitude or degree of difficulty to which an individual can rise. If we imagine all the tasks in the world arranged in order of difficulty, then the higher up this scale of difficulty we can climb, the higher would be our level of intelligence. The range of intellect would be the number of tasks at each degree of difficulty that we could solve. This range, of course, is determined by our opportunities for learning, because we cannot do many things unless we have had the chance to learn, but whether we are able to learn a given task or not will depend upon the level of our intelligence. In all intelligence testing, therefore, we must keep within the range of things common to all individuals that are to be compared.

For example, if we wish to compare in intelligence a child who has been to school with one who has not had such opportunity, we must not choose a test such as the National Intelligence Test which assumes a range such as is ordinarily experienced by school children only. Rather we must choose a test, like the Pintner Non-

Language Test that falls within the range common to both school and non-school experience.

Definition of Intelligence. — This discussion of intelligence has enlarged our concept of the word and given us a fair idea of what the psychologist at present means by it. Ordinarily when asked for a brief definition, he will answer by saying that intelligence is the ability of the organism to adjust adequately; or else he will stress the factor of modifiability and say, "Intelligence is the ability to learn, or to improve through experience." None of these definitions seems adequate, but that is of minor importance if we are cognizant of what is meant by means of the fuller description we have given above.

Summary

1. Our ability to judge intelligence without objective tests is very poor.
2. Experiments on judging the intelligence of children from their photographs show practically no agreement between the rankings of the judges and the test ranking.
3. Intelligence tests may be thought of as standardized interviews.
4. Intelligence tests may be divided into group and individual tests.
5. The most commonly used individual tests are revisions of the Binet-Simon Tests.
6. These tests are a miscellaneous group of tasks of all types arranged in levels of difficulty.
7. A mental age is obtained by comparing a given child's performance with the average performance of unselected samplings of children of each chronological age.
8. The intelligence quotient is the ratio of mental age to the chronological age, and it gives us a measure of brightness or dullness.
9. Performance tests are a second type of individual test for the measurement of intelligence. They may be said to measure concrete rather than abstract intelligence.

10. Group tests permit of less adjustment to each individual child than is the case with individual tests, but they are for that very reason well adapted for predicting success in school.
11. There are adequate group tests for every level of intelligence from the primary to the college level, as well as non-language tests and survey tests for special purposes.
12. It is important to remember that each type of test is testing a different aspect of intelligence and, therefore, we must not expect a child to obtain the same I.Q. on all tests.
13. Tests for pre-school children are now being rapidly developed and we are able to obtain a rough measure of intelligence even at 12 or 18 months.
14. In addition to tests of general intelligence there are several tests for special capacities, such as musical and mechanical ability.
15. The distribution of intelligence is normal.
16. The growth of intelligence is rapid at first, gradually decreasing as the child grows older.
17. The limit of growth has not been determined, but it is probably somewhere in the later adolescent period.
18. In the average child the growth is more or less constant from year to year, but in some children marked fluctuations seem to occur.
19. The effect of practice or coaching on intelligence tests is to increase the score. Practiced or coached children cannot, therefore, be compared with the norms based on unpracticed children.
20. Intelligence may be thought of as divided into different kinds, such as abstract, concrete and social.
21. Intelligence may be thought of as having different aspects such as speed, level and range.
22. Intelligence may be defined as the ability to adjust or the ability to learn.

Review

True-False Statements

As a review mark the following statements true or false:

1. Intelligence testing had its origin in the effort to separate the feebleminded from the normal.
2. If a child's I.Q. changes from 120 to 130 from one year to the next, then either the first or second test must be wrong.
3. Binet's first testing was done in order to discover superior children.
4. Three kinds of intelligence have been called abstract, concrete and social.
5. It is impossible to secure a valid measure of intelligence unless the child can read.
6. One can generally discover a person of low intelligence by his facial appearance.
7. The Stanford-Binet, the Pintner-Paterson Performance and the National Intelligence Test all measure identical abilities.
8. The intelligence quotient is obtained by dividing the chronological age by the mental age.
9. When we say that the distribution of intelligence is a normal distribution, we mean that there is an equal number of feebleminded, average and bright children.
10. The Terman Group Test is a good test for high school children.
11. The curve of growth of intelligence usually shows a sharp rise at adolescence.
12. It is possible to raise a child's score on the Binet by coaching.
13. An intelligence test is an adequate measure only as long as the individuals who take it have approximately the same background.
14. In testing foreign or deaf children, one should use a nonlanguage test.
15. We can measure a great growth of intelligence after the period of adolescence.
16. We cannot at present register any appreciable growth of intelligence after adolescence.
17. Mental development usually reaches maturity at the age of thirty years.

18. With children who develop normally the I.Q. increases with age.
19. Intelligence ratings are always comparisons with those who have been used as the standardization group.
20. Write in the three kinds of intelligence that have been suggested: (1) ————, (2) ————, (3) ————.
21. Underline the best response:
    The Binet-Simon Scale was first introduced into this country by: Terman, Goddard, Yerkes, Thorndike.
22. Before the name of each test below, put K if suitable for the kindergarten, E for the elementary school and H for the high school.
    ——— Terman Group Test
    ——— Pintner-Cunningham
    ——— Miller Mental Ability
    ——— National Intelligence Test
    ——— Pintner Rapid Survey
    ——— McCall Multi-mental
23. Useful revisions of the Binet Scale for the United States have been made by:
    1. ————
    2. ————
    3. ————
24. The intelligence quotient is obtained by dividing the ——— by the ———.

Advised Readings

Pintner, R.: *Intelligence Testing*, New York, 1923, Chapters 1–6.

Freeman, F. N.: *Mental Tests*, Boston, 1926, Chapters 1, 2, 4, 5, 6, 7, 13, 17 and 18.

References

1. Graves, K. B.: *The Influence of Specialized Training on Tests of General Intelligence*, Teachers College, Columbia University, N. Y., 1924.
2. Kuhlmann, F.: *A Handbook of Mental Tests*, Baltimore, 1922.
3. Pintner, R.: "Intelligence as Estimated from Photographs," *Psychological Review*, Volume 25, 1918, pages 286–296.

4. Pintner, R., and Paterson, D. G.: *A Scale of Performance Tests*, New York, 1917.
5. Seashore, C. E.: *Psychology of Musical Talent*, New York, 1919.
6. Stutsman, R.: "Performance Tests for Children of Pre-School Age," *Genetic Psychology Monographs*, Volume 1, Number 1, January, 1926.
7. Terman, L. M.: *The Measurement of Intelligence*, Boston, 1916.
8. Thorndike, E. L.: *The Measurement of Intelligence*, Teachers College Bureau of Publications, New York.

Group Intelligence Tests Mentioned in This Chapter

Pintner-Cunningham Primary Test: World Book Co., Yonkers, N. Y.

National Intelligence Test: World Book Co., Yonkers, N. Y.

Terman Group Test of Mental Ability: World Book Co., Yonkers, N. Y.

Thorndike Intelligence Test: T. C. Bureau of Publications, N. Y.

Pintner Rapid Survey Test: T. C. Bureau of Publications, N. Y.

Pintner Non-Language Test: College Book Co., Columbus, Ohio.

Otis Higher Self Administering Test: World Book Co., Yonkers, N. Y.

Army Alpha Test: Stoelting Co., Chicago, Ill.

CHAPTER VI

THE USES AND RESULTS OF INTELLIGENCE TESTING

Numerous Uses. — The uses to which intelligence tests have been put are manifold, and a detailed mention of them all is more suited to a special manual on intelligence testing than to an educational psychology. We shall, therefore, indicate in broad outlines some of their uses in the educational field. They are rapidly becoming an essential part of the organization of the modern school, and as such they represent one of the most useful contributions of modern psychology to the practical business of teaching.

Classification. — The most common use of intelligence tests, group tests especially, is as an aid toward better classification of pupils in grades or in the various sections of a grade. Children differ greatly in general intelligence and in the rate of growth of their intelligence. They differ in mental age or the present amount of intelligence possessed by them, and also in their intelligence quotients or the rate at which their intelligence is probably growing. In most schools where little attention has been paid to homogeneous classification the differences in intelligence among the children of any one grade are very great.

In Grade VB in a certain school in New York the National Intelligence Tests were given and the frequency distribution according to mental age of the three sections of this grade was as follows:

MENTAL AGE (N.I.T.)

Grade	*7–0 to 7–11*	*8–0 to 8–11*	*9–0 to 9–11*	*10–0 to 10–11*	*11–0 to 11–11*	*12–0 to 12–11*	*13–0 to 13–11*	*14–0 to 14–11*	*15–0 to 15–11*	*Med. M.A.*	*Total Cases*
5B1		3	8	13	13	1	3	0	1	10–7	42
5B2	1	3	8	11	8	5	5	0	1	10–9	42
5B3	2	3	12	14	2	2	1	0	0	10–0	36

This means that of the forty-two children in the first section of the grade three have mental ages between eight years no months and eight years eleven months; eight have mental ages of 9; 13 have mental ages of 10; and so on. The median mental age of that section is ten years seven months. In these three sections there are children with mental ages from seven to fifteen, presumably all doing the same type of work. Undoubtedly there are many misfits, many who are really unable to do the work, and many who are capable of doing much more advanced work. Those at the lower end of the distribution have probably been promoted more or less automatically and very likely are totally unfitted to do the work of the fifth grade. Those at the upper end might well profit by skipping a half grade or by an extra promotion. Consultation with the teacher about the child with reference to his school work, health and so on, is of course assumed, before such changes are made. The mental age, therefore, helps us to determine the grade or level at which the child might best be working.

Mental Age to Determine Grade. — In the specific example we have cited above, the teacher or principal would begin by finding out whether those with the lowest mental ages of seven and eight could really do the work of the class. As this is very unlikely, they would either be put into some special class or held in the same class at the end of the term, or, as a last resort, demoted. Many of the nine-year-olds should be dealt with in the same way, but as these form a goodly number of the total class, the chances are that the instruction of the group has been more or less adjusted to their level, and, therefore, many of them would remain. Those having mental ages of 10, 11 and 12 would be left in the class and they would be expected to do normal work. The nine children with mental ages of 13 would probably profit by being promoted at once to the next higher half grade, regardless of their actual educational performance in the present class. They would be expected to do better work, and in most cases they would come up to expectation. Lastly, the two children with mental ages of 15 would be skipped a half grade or be given two promotions each half year until they arrived at a grade more suitable to their mental level. In some such way as this the great spread of mental ages in one grade, due to promotion by chronological age, can be gradually restricted. The wise superintendent or principal will not attempt a sudden revolution in grade placement, but will in the way suggested above work gradually from year to year.

The I.Q. for Sectioning. — Assuming, then, that such necessary shifts from one grade to another have been made and that the remaining children are probably all able to do the work of the grade, then the I.Q., which tells us the rate of growth of intelligence, may be used to help us make a division of the children into different sections so that they may move along at different rates of speed. At the same time the brighter sections will probably do more extensive work than the slower sections. Broadening the field of study and acceleration of pace generally go together. Here is the distribution of the I.Q.'s of a VA Grade tested on the National Intelligence Tests: —

I.Q.	*Frequency*
130–139	2
120–129	2
110–119	14
100–109	16
90–99	22
80–89	18
70–79	11
60–69	7
50–59	1
Total	93

For reasons of space these 93 children had to be divided into three equal sections. Instead of having a random selection of all I.Q.'s in each of the three sections as was actually the case, it would be better to have a section of 34 with I.Q.'s above 100; another of 31 with I.Q.'s from 85 to 100 and a third of 28 with I. Q.'s below 85. No hard and fast lines should be laid down and shifts from one section to another should be readily made for any good reason.

The number of children in a grade, the number of classrooms and teachers available, and other actual administrative conditions will determine in each case the number of sections that can be formed. The actual intelligence of any particular school population will determine where the lines should be drawn between one section and another. What might be considered an advanced or bright section in one school will very likely be considered quite ordinary or indeed mediocre in a school having a highly intelligent population. There are schools in which the average I.Q. is barely 90 and others in which the average I.Q. is 115. The extreme form of the recognition of individual differences in intelligence is found in those schools which attempt to let each child go at his own rate. There are several well-known "plans" of school administration which might serve as examples of attempts to recognize and provide for each individual's different intellectual needs.

Example of Use of M.A. for Classification. — School classification by mental age as suggested above has been reported by many investigators. One such report (Pintner and Noble, 7) deals with a school of 450 children in grades I to V inclusive. All the children were tested with intelligence tests. Allowing one and one half mental years as normal for each grade, it was found that 56.7% were normally placed according to mental age; 13.1% were above normal and 30.2% below. A reclassification was made in consultation with the teachers. Sixty-nine children were given extra promotions, 41 were not promoted at the mid-year and 7 were demoted. A small ungraded class for very low I.Q.'s was formed. This reorganization worked well. At the end of the next semester the percentage of normally placed pupils was 71%; below normal 16.2%; above normal 12.7%. Reports from the assistant superintendent, principal and teachers all expressed satisfaction in the results achieved, not only with respect to increased teaching efficiency, but also with respect to the general morale of the school.

Under this general heading of classification we may say that intelligence tests help in:

(a) grading or placing the child in the correct grade.
(b) sectioning or placing the child in the correct section of a grade, such as fast, medium, or slow sections.
(c) acceleration or deciding whether a child should be given extra promotions or should skip a grade and thus accelerate his general progress through the grades.

Selection of Dull and Feebleminded. — One of the earliest uses of intelligence tests in school was as a means for the selection of the very dull and feebleminded so that they might be taught in special classes. Such special classes were in existence before the coming of the intelligence test, but the latter has been of great help in the selection of such children.

Assignment of children to a special class should not be made on the basis of a group test. A thorough individual examination by a competent person is essential, because the practical consequences of such segregation are very serious for a child. Furthermore children in such special classes should be re-tested from time to time in order to verify the first mental rating and to see if some improvement may not have taken place. This is unfortunately rarely the case.

The intelligence quotients of children in such classes generally range from about 50 to about 70. Few children with I.Q.'s below 50 stay long in the public schools. They are generally sent to some institution for the feeble-minded, because they need a type of training and care which the public school can hardly give. Those above 70 are generally able to succeed in the ordinary classroom, although most special classes have some cases of I.Q.'s in the seventies and eighties. It is bad policy, however, to put any and all misfits into the same special class. Disciplinary cases with I.Q.'s above 80 or 90 do not belong in the same class with those of low I.Q.

The distribution of the I.Q.'s of the special class children of one large New York public school is shown below (Irwin and Marks, 4):

I.Q.		*n*	*Per Cent*
71–75		31	20.7
66–70		45	30.0
61–65		31	20.7
56–60		23	15.3
51–55		16	10.7
46–50		3	2.0
41–45		1	0.6
	Total	150	100.0

In this distribution we note that the policy of the school is not to admit any with an I.Q. above 75, but to include a fair number between 71 and 75. Again we note the very small number of cases with I.Q.'s below 50.

Selection of Bright or Gifted. — It was the coming of telligence tests that really discovered the child of superior intelligence, in the sense of the well-rounded individual of great potentialities. The number of children of high I.Q.'s above 130 or 140, is probably about 1% of the total school population. Several schools are experimenting with special classes for such children. If brought together in a special class they have more opportunity to expand and develop. Certain difficulties inherent in a policy of mere acceleration are avoided. Their influence upon each other is generally very beneficial. As a rule such children are taller, heavier, healthier than the child of average I.Q. They are just as much interested in games, they love to play just as much as the average child. They generally read a great deal more extensively than does the average child. They are not generally aware of their superiority in intelligence and in the ordinary class many are often found doing merely average work. From the first year in school they have found such work commended. They have satisfied their teacher. They do not know that they are capable of much higher achievement, because they have never been stimulated to do more. Many, therefore, form habits of intellectual work much below their real capacities and these habits may continue throughout their whole educational career. These are the kind of people who sometimes do exceptionally well in some great emergency or when suddenly called upon to do a task out of their usual line of work. They surpass themselves, and we remark, "I did not know he had it in him." He did not know he had it in him either, because he never had received an adequate stimulus that could call forth all his capacity. Such people do not surpass themselves, they merely surpass the standard that the school set for them and which they were willing to accept. It is very important, therefore, that bright children should be recognized as early as possible in their school career, before they have had time to acquire habits of mental work below their level. Intelligence testing should begin as soon as the child enters school. It is highly important for society that these intellectual resources should be conserved and not wasted through ignorance and neglect.

It has been shown that these very bright children respond readily to an accelerated and broadened curriculum.

They do not need the same amount of drill in the fundamental subjects. They can initiate and plan much better than those of average intelligence. When grouped together in the same class they serve as stimuli to each other and the competition of keen minds prevents the rise of snobbishness or vanity. Just what is the best course of study for such children we do not yet know. That it is not the course of study now adapted to the average child is obvious. Much experimental work will be necessary to find the proper material and the proper methods for these active minds.

There is no sharp dividing line between normal and bright children. In the formation of classes for bright children the lower limit in terms of I.Q. has been set anywhere from 110 to 155 according to the number of children available. Terman's (9) extensive study was based on children with I.Q.'s above 135. The special conditions existing in any particular school will determine what children should be segregated for special instruction. In some schools having an average I.Q. of 90 or 95, it might be feasible to form special classes for those with I.Q.'s above 110. In other schools having an average I.Q. of 110 or 115, such a limit would include fifty per cent of the children, and therefore more highly selected groups with I.Q.'s above 130 or 140 might be formed. In any large school the class or classes for bright children should be made up of those possessing the highest I.Q.'s within that school.

Guidance. — In a broad sense all intelligence testing may be used for general educational and vocational guidance. In the narrower sense of giving specific educational advice as to whether to go to high school or college, or as to the course of study to be chosen, intelligence tests may help but they cannot be used in a mechanical or arbitrary way. The wise counselor will always make use of them but never depend upon them alone. Achievement in school or college and success in later life do not depend solely upon intelligence. Character qualities are very important as well as temperament, health and opportunity. Intelligence, however, prescribes the limits within which the individual must function.

Abstract intelligence, as tested by most of our verbal intelligence tests, is required for most of the regular school subjects. Unless, therefore, a child possesses a certain amount of it, he is not likely to be successful in mastering them. Thus a child with an I.Q. of 80 is unlikely to be able to complete the eighth grade successfully. Such a child ought to be given practical vocational training in the last few years of his school life so that at the age of 14 he may be in some way prepared for his future work in the world. A child with an I.Q. below 90 will probably derive little profit from the ordinary high school course. Children with I.Q.'s between 90 and 100 will probably profit from some of the work in high school, but few are likely to graduate without difficulty. In almost all cases with I.Q.'s above 100, the children will be able to take the regular course, although difficulty may be found in some of the more abstract work. To grasp the symbolism, understand the proofs and make the generalizations required in algebra probably requires a minimum I.Q. of 110. It is doubtful whether children with I.Q.'s below 110 should be advised to take the customary course at a liberal arts college.

College Entrance. — A great many colleges make use of intelligence tests for all entering students and some use the results of such tests in helping to determine who shall be admitted. No college uses the intelligence test alone in making such a decision.

During the past decade college education has become extremely popular and great numbers of high school graduates are seeking admission to the college or university. In the state of Indiana an intelligence survey

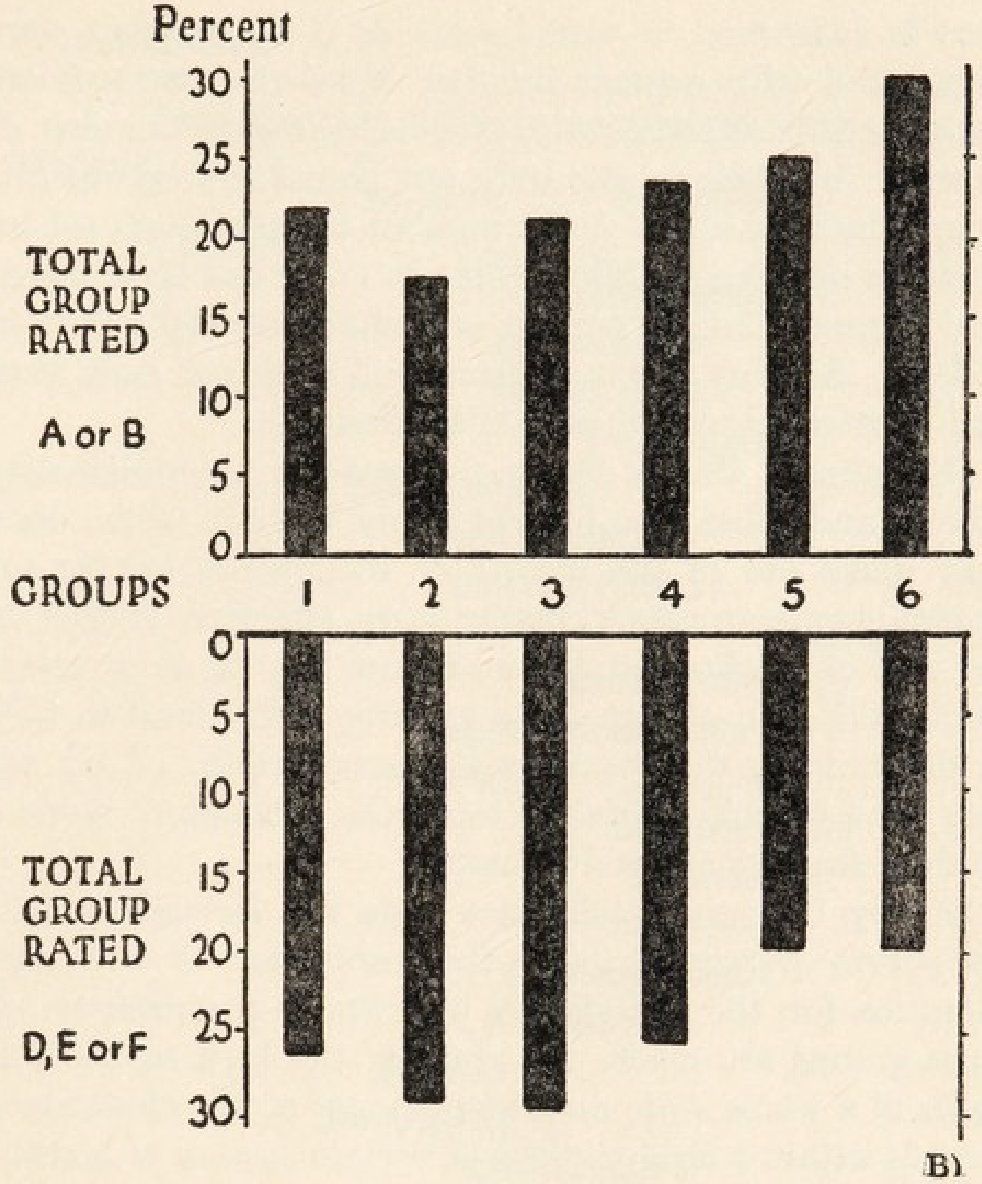


Figure 30. — Percentage of high school seniors possessing the highest (A or B) and the lowest (D, E or F) grades of intelligence in the following groups: (1) our total or standard group; (2) all who will not attend college; (3) those expecting to attend college, no college selected; (4) all who expect to attend college; (5) those expecting to attend a technical or professional school; (6) those expecting to attend a college of liberal arts. (From Book's *The Intelligence of High School Seniors.*)

(Book, 1) of high school seniors showed that those expecting to attend college were very much like the total group of seniors tested. Figure 30 shows the percentages of students with various college intentions as compared with the total group. So far as the intention of the student is concerned, it would seem as if the colleges were threatened with a great number of inferior or mediocre individuals, and conversely were losing a great number of superior intelligence who were not planning to go to college. Undoubtedly a great number of those high school students of inferior ability, who do carry out their intention of proceeding to college, will encounter difficulties in college. A great many of them will drop out each year and relatively few will remain to graduate.

Disciplinary Use of Tests. — Teachers, administrative officers and others are finding many ways in which they may make use of the results of intelligence testing for disciplinary purposes. Particularly effective is this in the case of brighter students who are lazy or doing inefficient work. In college, tests are frequently used to help in determining dismissal for low scholarship. They are also used to help in determining the amount of work a student shall be allowed to carry.

Survey Uses. — Intelligence tests are frequently used for survey purposes, where the results are of little significance for the individual, but where comparisons of large groups are made. Surveys of the high school students of a whole state may be made, or of the elementary schools within a city or within a certain section of a state. A survey of a university may be made in order to compare the intelligence of the students in the different colleges of the university. The largest survey by means of intelligence tests was that made during the World War of more than a million drafted men in the United States Army. A sample of the type of comparison of schools in a survey is shown in Figure 31, which shows the middle fifty per cent of high school seniors in various schools in various counties in Indiana. The great differences between schools are well brought out.

Measures of Progress and Efficiency. — Intelligence tests are of importance in helping to measure the educa-

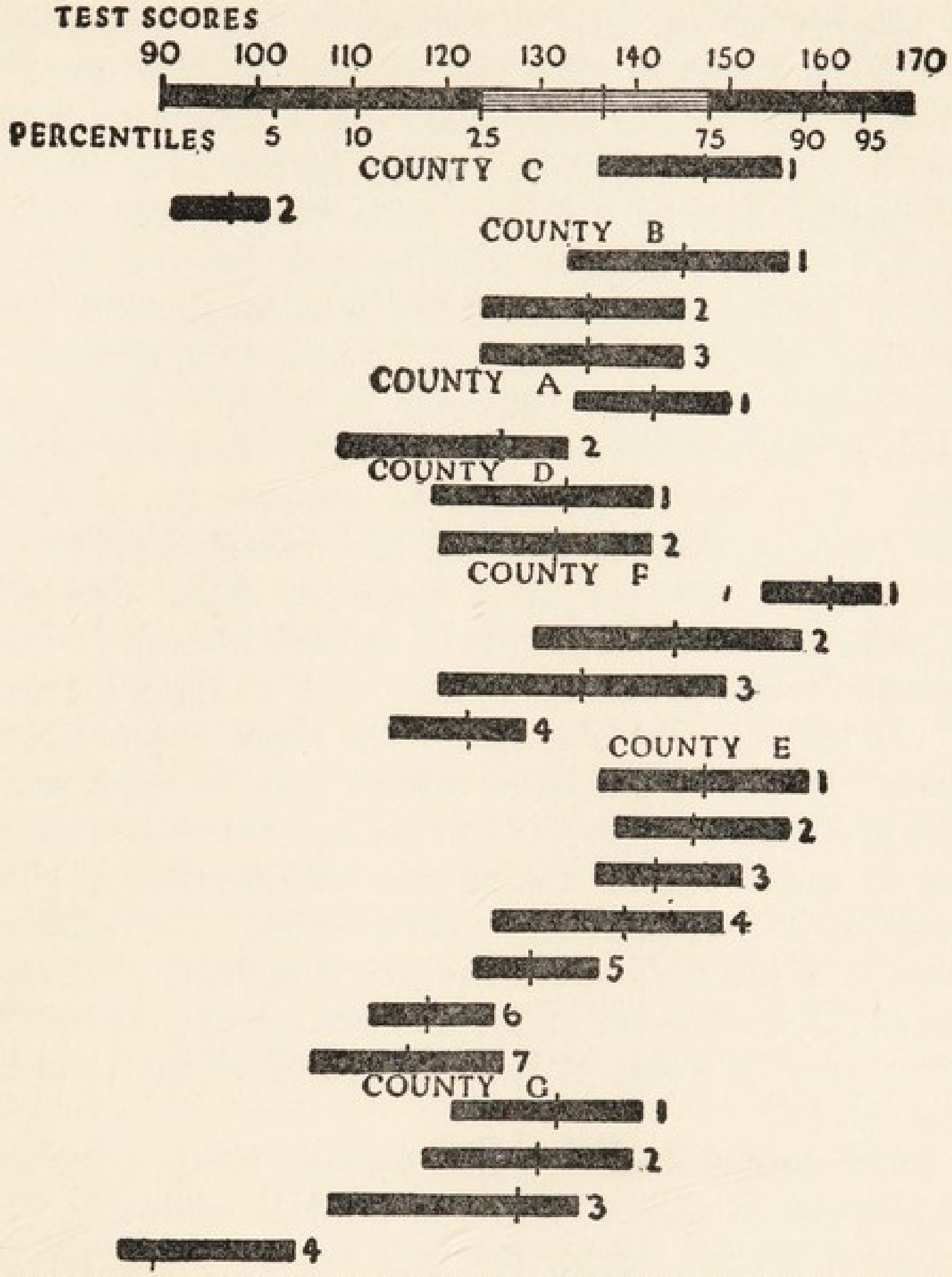


FIGURE 31. — Scores made by the middle 50 per cent of seniors representing individual schools in 7 representative counties selected at random from the 92 counties in the state. Figures at right of bars designate the individual schools. (From Book's *The Intelligence of High School Seniors.*)

tional progress of pupils on the one hand and the efficiency of the teaching on the other. This is done by means of a comparison between a child's intelligence and his school achievement. Obviously the child of low intelligence cannot be expected to make the same amount of progress as the child of average or high intelligence. We must therefore take the intelligence into account when evaluating any individual's school progress. What may be considered splendid progress for a child of average intelligence may be very poor for a child of high intelligence. To rest content with average progress in the case of children of superior ability is to countenance idleness and inefficiency. It means a waste of intelligence.

A teacher's efficiency can only be justly evaluated in terms of the intellectual material with which she has to deal. It is manifestly unfair to compare the work of two teachers without knowing the mental make-up of their pupils. And yet this is frequently done. One who is achieving excellent results is praised without further question, whereas another teacher with an inferior group of children may be blamed for the small improvement made. In reality the latter teacher may be much more able and efficient. No real comparisons between teachers can be made without taking into consideration the intelligence of the children they are teaching. The measurement of a teacher's worth lies in the amount of change she can produce in her pupils, and this can only be rightly interpreted in terms of the intelligence of the children themselves.

Experimental Uses. — Lastly we may mention the use of intelligence tests in scientific work in education. In countless studies during the past twenty years, intelligence tests have found a recognized place. One of the most common ways in which they are used is to help the experimenter in the formation of equivalent groups. We wish to measure the effect of a new procedure, a new method. We raise such questions as to which of two methods of reading is best, as to whether the study of Latin increases our knowledge of English words, whether dental hygiene improves scholarship, and the like. If our method is to compare an experimental with a control group, then we must make our two groups as nearly alike as possible in all ways, and one of the most important ways in which two such groups must be kept alike is undoubtedly with respect to intelligence. If they are not so equated, the increase in efficiency of the experimental group may not have been due to the method of teaching arithmetic employed, or the fact that they have had their tonsils removed, but simply and solely to the fact that their I.Q.'s are higher and they have therefore been developing intellectually faster than the control group.

Intelligence and Occupation. — Different occupations obviously require different amounts of intelligence. Just how much is required for each occupation we do not know, but we may get some idea by studying the intelligence of those now holding different positions. In the rough and ready selection of everyday life, people will tend to drop out of those occupations for which they do not have the requisite amount of intelligence.

The most comprehensive survey of the intelligence of various occupational groups was made in connection with the army testing. Figure 32 shows the intelligence ratings of the middle fifty per cent of each occupation, the short vertical line marks the median intelligence and the total scale of intelligence is pictured at the top and bottom of the diagram. The median intelligence of laborers, for example, is C−, i.e. rather mediocre intelligence. The middle fifty per cent stretch from D, through C− to C. In addition there was a lower group of twenty-five per cent who made ratings of D and D−, and an upper group also of twenty-five per cent who made ratings of C, C+, B and A. As a matter of fact of the 1453 laborers tabulated, 10.7 per cent made ratings of C+, 3.5 per cent ratings of B and 0.62 per cent ratings of A. In a similar manner we may interpret all the other occupations on the chart.

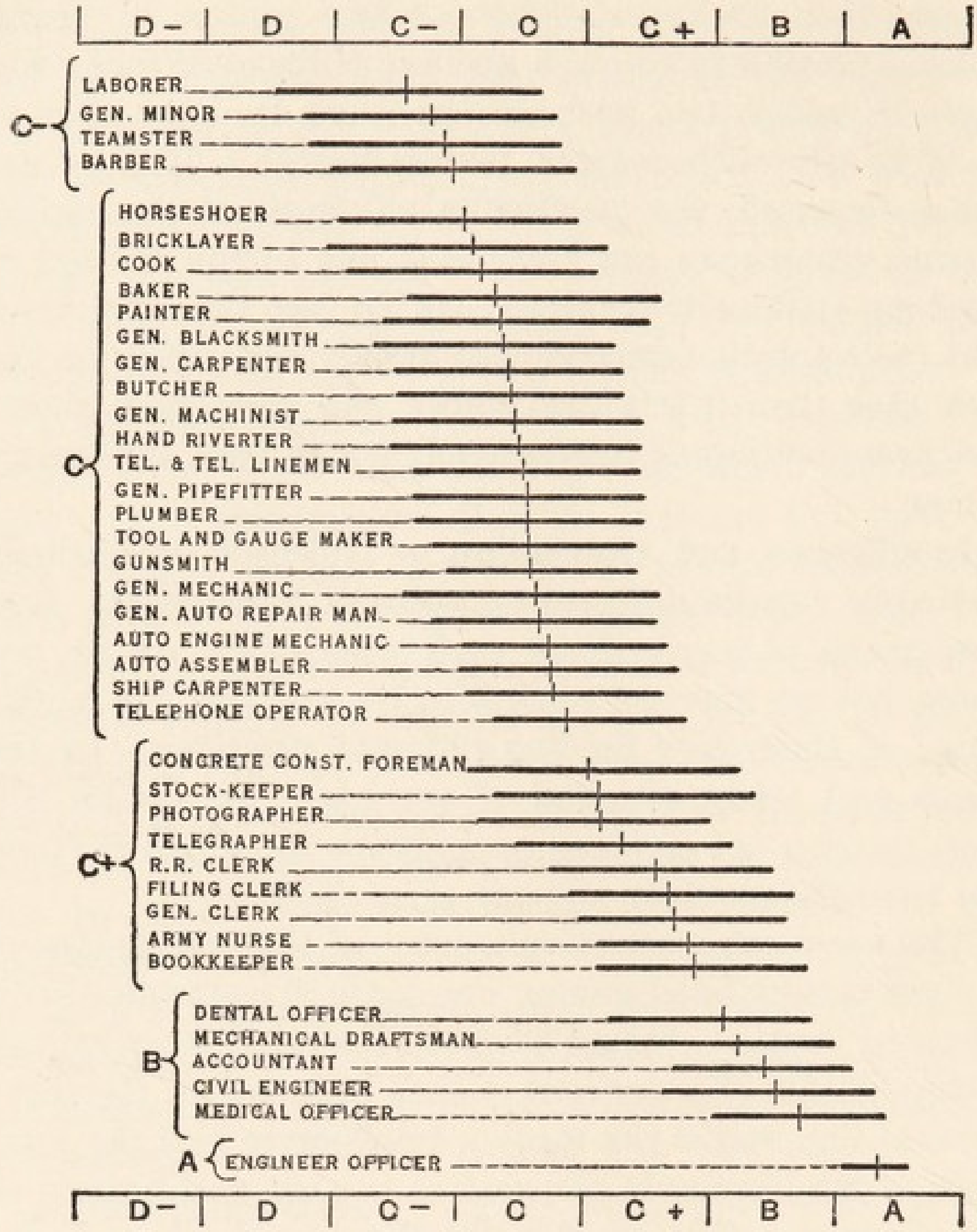


FIGURE 32. — Occupational intelligence ratings. Letter grades on horizontal scale. Length of bar for each occupation is mid-range of 50 per cent; median point is shown by a cross line. (From *Memoirs of National Academy of Sciences,* Vol. XV.)

The chart should not be considered as absolutely accurate for all occupations in ordinary peace-time conditions. The draft was somewhat selective. Nevertheless in its broad outlines it is undoubtedly significant. It shows how selection according to occupation is taking place all the time. From the point of view of general intelligence alone, it shows the wide selection of occupations open to a person of a given intelligence rating. The man of high intelligence has the whole range of occupations to choose from. The man of C + intelligence has a fairly wide selection, particularly if he is not ambitious to surpass the average of his group. If he wishes to belong to the upper twenty-five per cent of his chosen occupation, his choice is considerably narrowed. For the man of D intelligence, the choice is very narrow indeed.

Other rankings of occupational groups have been made but none on so vast a scale as the army chart. Table 1 shows the average score on an intelligence test given to various occupational groups (Scott and Clothier, 8). The

TABLE 1

AVERAGE SCORE OF TYPICAL GROUPS

Groups	*Average Score*
Sales force, employed to assist in holiday rush (women)	25
Sales force, department store (women)	27
Students, commercial business college (women)	28
Sales force, employed to assist in holiday rush (men)	29
Office boys	31
Sales force, department store (men)	33
Machine operators (men)	33
Job foremen	38
All office employees (women)	40
Foremen	41
Nurses	42
Rotary Club members	46
Executives of progressive firm	51
Supervisors (manufacturing plant)	52
Sales executives	54
Students (college of arts and sciences)	54
Students (medical school)	56
Engineering students	57
College presidents (small colleges)	58
Students (succeeding in examinations for internship)	60

difference between the upper and the lower groups is quite marked.

The use of intelligence tests in a large tire manufacturing company led to a decision that individuals having scores below certain points were not good risks. Figure 33 shows the average intelligence scores of four different occupational groups in this factory and also the critical scores below which an applicant for a position should not fall without having definite compensating qualifications.

Tests of Specific Capacities. — In addition to such general tests, mention must also be made of tests for specific capacities. The best example at the present time is the Seashore Test of Musical Capacity, which measures such traits as pitch, intensity, time, rhythm, consonance and auditory memory. These tests predict musical ability and agree well with the estimates of music teachers as to their students' abilities. The Eastman School of Music of the University of Rochester requires that these tests be taken by all candidates for admission to the school. Those whose talent, as measured by the test, does not warrant the expenditure of time and money on the study of music are advised accordingly. It is obviously to the interest of both the student and the school to have such tests made.

These are only a few of the ways in which general intelligence tests and specific capacity tests may be used in the wide field of vocational guidance and industrial placement. Intelligence tests cannot place the right man in the right job, but they do help, and any help given in this important work of steering the individual to a fitting occupation is well worth while not only in terms of happiness and contentment to the individual, but in terms of the general efficiency and moral health of society in general.

Intelligence and Delinquency. — One of the interesting results of intelligence testing has been the study of the mentality of the delinquent individual. A great many studies have been made, and, although the percentage

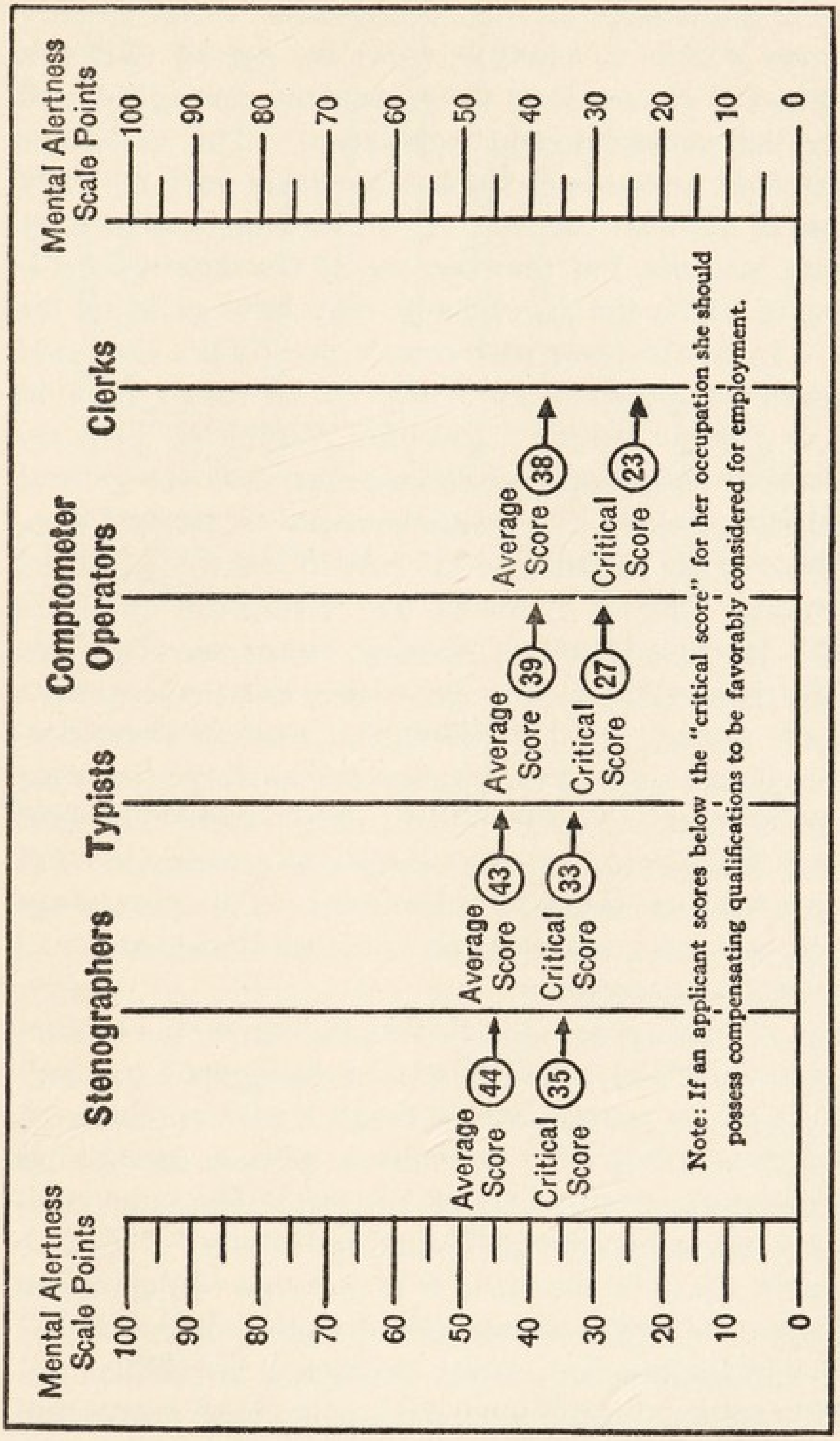


FIGURE 33. — Occupational Standards on Mental Alertness Test for women applicants and women employees in a large tire manufacturing company. (From Scott and Clothier's *Personnel Management.*)

of feeblemindedness among delinquents varies greatly from one worker to another, most are agreed that this percentage is greater than the percentage generally found among the non-delinquent population. The variations in estimated amounts of feeblemindedness vary not only because of different concepts of feeblemindedness held by different workers, but also because of the different institutions in which the psychologist may have gathered his cases. In workhouses and county jails more dull and feebleminded prisoners are likely to be found than in state or federal prisons. Juveniles committed to a reform school will on the whole be duller than the general run of those found in a juvenile court detention home, and these again be duller than those in the special school for truants. Figure 34 shows the intelligence quotients of 1212 juvenile repeated offenders tested by the Stanford Revision (Healy, 3). From this chart we see that there are a great many children of normal or above normal intelligence who are delinquent, but there is an excess of cases testing below 90 I.Q. From a study of 4000 repeated delinquents the same investigator concluded that 13.5 per cent are clearly feebleminded. This percentage of feeblemindedness is very much higher than that found in the general population.

Lack of intelligence is undoubtedly one of the numerous causal factors at the basis of delinquent conduct, but it is by no means the sole cause of such conduct. A careful comparison of a delinquent with a comparable non-delinquent group of children living in the same environment has been made by Burt (2). He says,* "The following proves to be the order of importance of the various conditions we have reviewed: (1) defective discipline; (2) specific instincts; (3) general emotional instability; (4) morbid emotional conditions, mild rather than grave, gen-

* Quoted by permission from Burt, C., *The Young Delinquent*, D. Appleton and Co., New York.

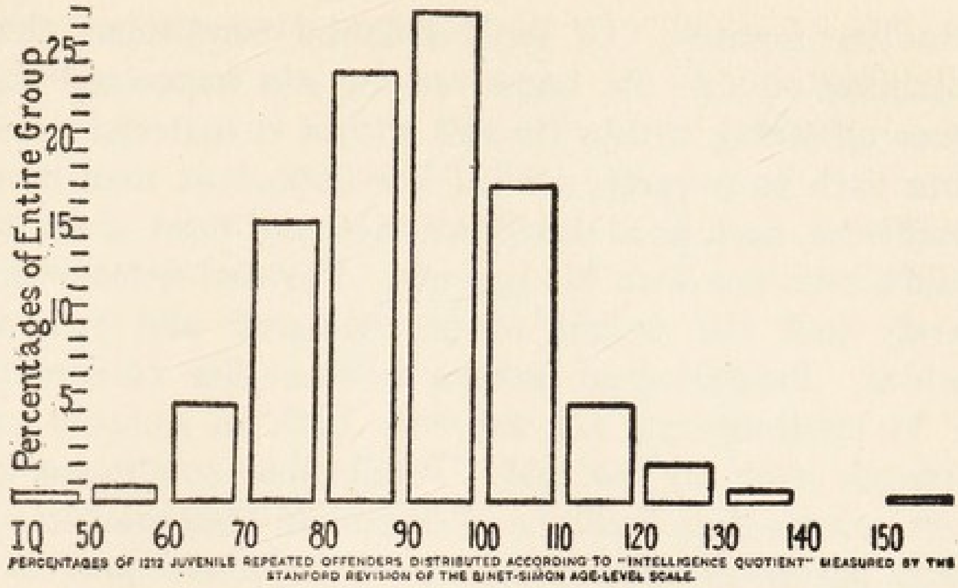


FIGURE 34. — Distribution of I.Q.'s of repeated offenders. (From Healy's *Practical Value of Scientific Study of Juvenile Delinquents.*)

erating or generated by so-called complexes; (5) a family history of vice or crime; (6) intellectual disabilities, such as backwardness or dullness; (7) detrimental interests, such as a passion for adventure, for the cinema, or for some particular person, together with a lack of any uplifting pursuits; (8) developmental conditions, such as adolescence or precocity in growth; (9) a family history of intellectual weakness; (10) defective family relationships — the absence of a father, the presence of a stepmother; (11) influences operating outside the home — as bad street companions and lack or excess of facilities for amusement; (12) a family history of temperamental disorder — of insanity or the like; (13) a family history of physical weakness; (14) poverty and its concomitants; and last of all, (15) physical infirmity or weakness in the child himself.

"Heredity appears to operate, not directly through the transmission of a criminal disposition as such, but rather indirectly, through such constitutional conditions as a dull or defective intelligence, an excitable and unbalanced temperament, or an over-development of some single primitive instinct. Of environmental conditions, those obtaining outside the home are far less important than those obtaining within it; and within it material conditions such as poverty, are far less important than moral conditions, such as ill discipline, vice, and most of all, the child's relations with his parents. Physical defects have barely half the weight of psychological and environmental. Psychological factors, whether due to heredity or to environment, are supreme both in number and strength over all the rest. Intellectual conditions are more serious than bodily, and emotional than intellectual; while psycho-analytic complexes everywhere provide a ready mechanism for the direction of overpowering instincts and of compressed emotional energy into open acts of crime."

The Deaf and the Blind. — The education of these children, deprived as they are of one or the other of the two most important senses, is for that reason very difficult. Intelligence tests would seem also to suggest that on the average they are somewhat below the hearing or seeing child. This is very marked in the case of the deaf. With verbal tests the difference is very great, with non-verbal or performance tests the difference is not so great, but still marked. This makes the education of such children a doubly difficult task. In the case of the deaf especially the lack of hearing closes the most important avenue for the learning of language and hence slows up and makes more difficult the process of learning to read. Thus those, who might hope to compensate for their deficiency somewhat through the printed page, are handicapped at the start by the great difficulty in the mastering of reading and also by a slighter endowment of general intelligence.

Figure 35 shows a comparison of 678 deaf fourteen-year-olds with 1120 hearing fourteen-year-olds on the Pintner Non-Language Test. The percentage of deaf

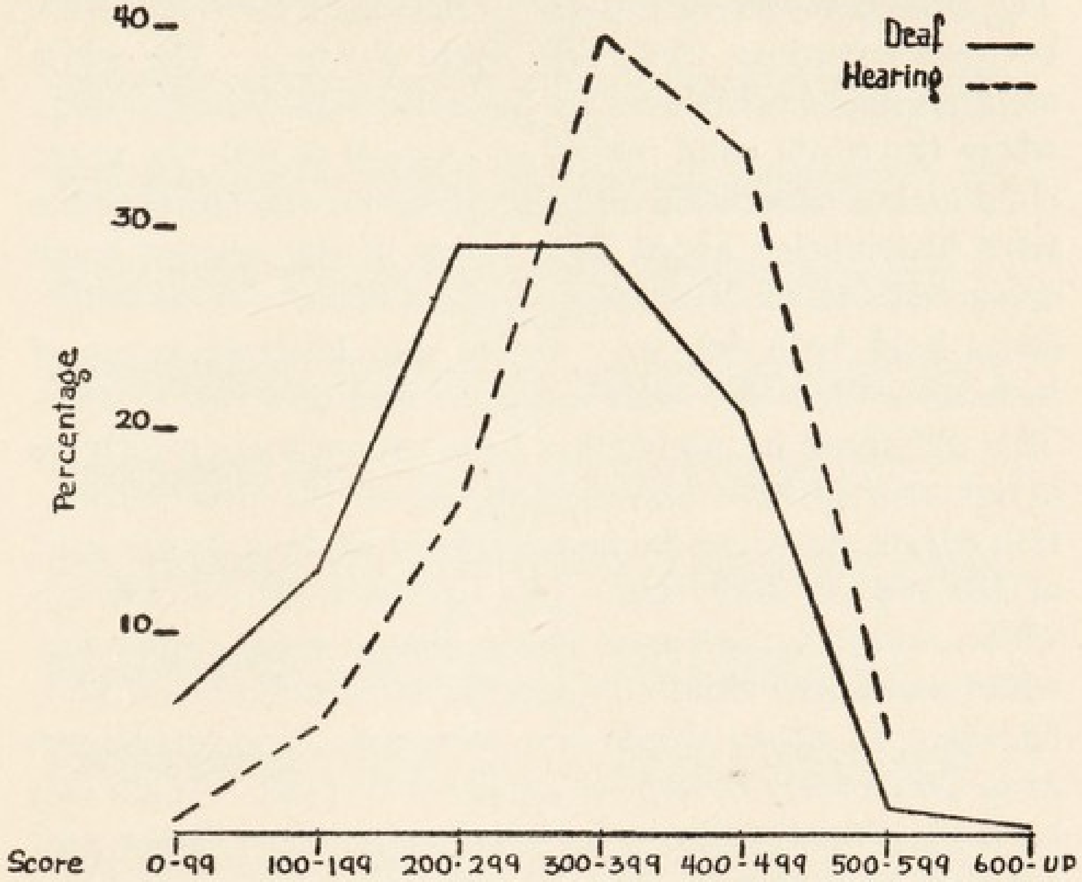


FIGURE 35. — Distribution of scores on the Pintner Non-Language Mental Test of Deaf and Hearing Children—Age 14. (From Day, Fusfeld and Pintner's *Survey of American Schools for the Deaf.*)

children making low scores is much greater than the percentage of hearing children, and, conversely, the percentage of deaf making high scores is lower than the equivalent percentages for the hearing. The mean score of the deaf fourteen-year-olds is about the same as the mean score for eleven-year-old hearing children (Pintner, 5).

Intelligence and Race Differences. — The interest of man in his fellow men of different racial origin leads him continually to make comparisons and contrasts. It is only natural that many comparisons of different racial groups by means of intelligence tests should have been made. For the most part these comparisons have been of school children of different racial stocks in this country.

The most extensive comparisons that have been made are between American white and negro children. The white children make a higher average score. In terms of I.Q., where the white child makes an average of 100, the negro child makes an average of from 85 to 95. In terms of the total distribution about 25 per cent of the negroes reach or exceed the median score of the whites. These differences hold both for very verbal and abstract types of tests as well as for more concrete and non-verbal tests. This difference in intelligence between negroes and whites is not very great. The overlapping of the two distribution curves shows, as we have said above, that 25 per cent of the negroes have more intelligence than the average white, and it furthermore shows that a great number of white and negro children have the same amount of intelligence. In other words the two racial groups do not form two widely divergent intelligence types. There can be no justification from the standpoint of intelligence of a policy of separate education for the two groups. If such is deemed advisable it must be justified on other grounds. Nor can race prejudice against the negro find support in these test findings. Race prejudice springs from and is nurtured by emotional reactions and is not occasioned by scientific findings, although we frequently make use of such findings to rationalize our prejudices. To say that you discriminate against the negro because he belongs to a race having inferior intelligence, is an attempt on your part to rationalize your racial prejudice.

Other racial groups in the United States have been compared by means of the results found among drafted men in the army testing. A great many men of foreign birth were found in the draft army. The distribution of the intelligence grades for many national groups is shown in Figure 36. The white shows A or B (i.e. superior ratings), the shaded shows C or average ratings and the black D or inferior ratings. These ratings, of course, tell

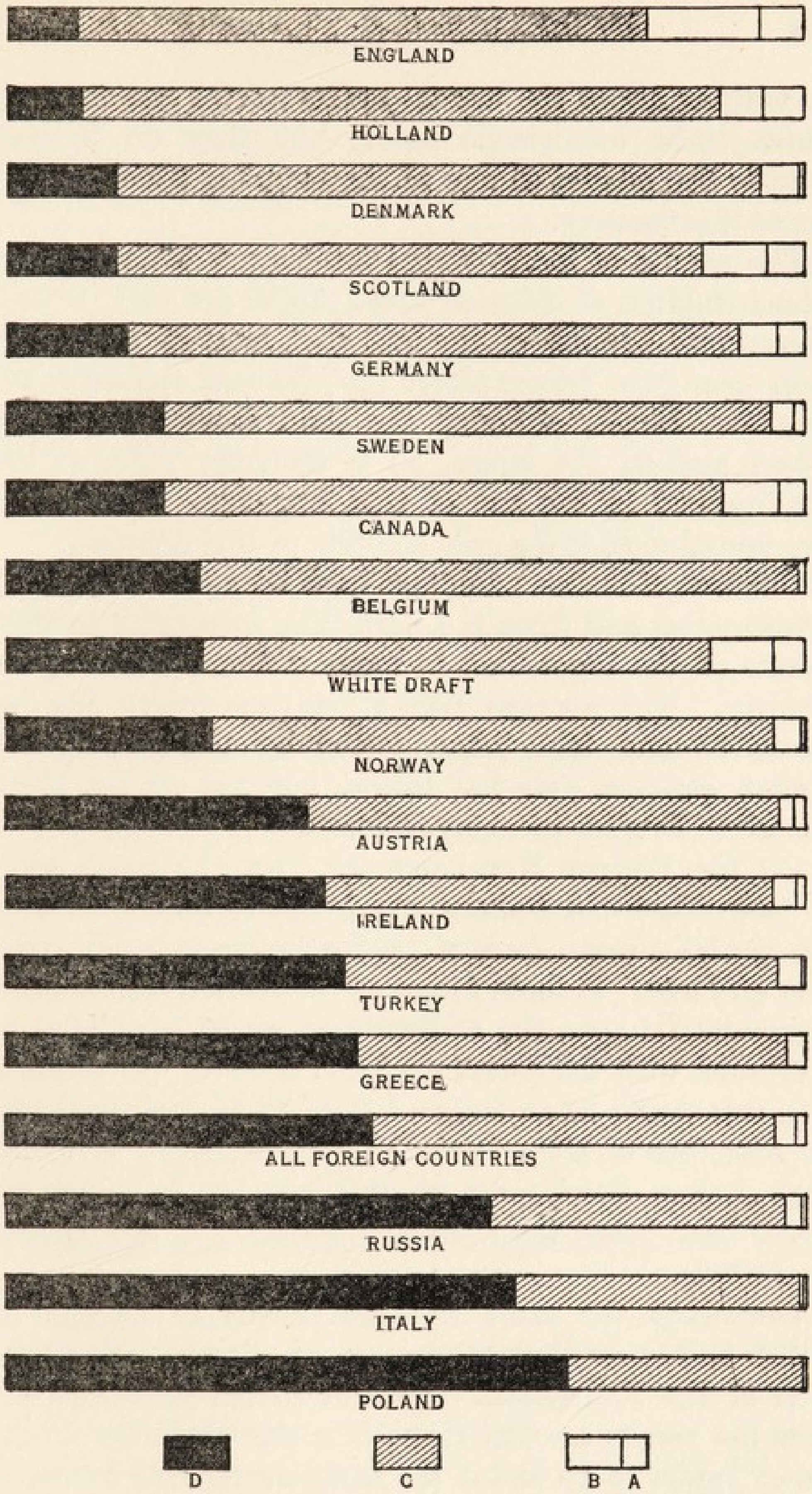


FIGURE 36. — Percentage distribution of letter grades in intelligence by nativity of foreign-born men in army draft. (From *Memoirs of the National Academy of Sciences*, Vol. XV.)

us nothing about the mental ability of the nations from which these immigrants come, but they do indicate roughly the mental calibre of the sampling which is coming to this country.

The results of many verbal intelligence tests given to school children of different racial stocks are very difficult to evaluate because of the fact that most of these children come from homes where the language spoken is not English. Hence, although the child speaks English in school and on the street, he is probably more or less handicapped in the use of this language. The use of non-verbal tests is the only solution of this question.

Children of Italian parentage have been tested by many investigators and there is a surprising unanimity in their results on the Binet Test in many different parts of the country. The average I.Q. of various groups lies between 85 and 90. When tested by non-verbal tests Italian children also fall below, but not so far below American English-speaking children.

On the Pintner Non-Language Test the mean score of twelve-year-old Italians was found to be 305 as compared with a mean score of 329 for twelve-year-old American children. If these scores are converted into approximate mental ages, the Italians have an M.A. of 11–6, as compared with the American M.A. of 12–6. In terms of I.Q. this would be an Italian I.Q. of 92 as compared with an American of 100. It would seem probable, therefore, that Italian immigrants to this country have on the whole been below the average intelligence of the country as a whole.

Somewhat the same situation seems to be true of Polish and other Slavonic races, with the probable exception of the Bohemians. With reference to the Jewish race the results are not clear. On the whole the intelligence seems to be about the same as that of American children, with a possibility of being slightly superior.

American Indian children fall below the average on tests of abstract intelligence.

All of these results of the testing of children of foreign parentage give us merely a clue to the intelligence of the immigrants from foreign countries. They cannot in any sense give us any knowledge as to the average general intelligence of the countries from which these immigrants have come. Up to the present time no adequate comparative tests have been made between peoples in the countries where they live. One very inadequate sampling of Belgian children made in Belgium showed no difference between their scores and the American norms on the Non-Language Test (Pintner, 6).

Summary

1. The most common use of group intelligence tests is an aid to the classification of pupils.
2. The mental age helps to determine the grade for which the pupil is best fitted.
3. The intelligence quotient helps to determine the section for which the pupil is best fitted.
4. Intelligence tests are used to select the dull and feeble-minded for special instruction.
5. Intelligence tests are essential for the selection of bright children for special classes.
6. It is very desirable that such children should be selected as early as possible in their school careers.
7. In a negative sense the results of intelligence tests are useful for general educational advice and guidance.
8. Many colleges are using intelligence tests as an aid in the selection of their students.
9. Surveys of high school students show that many of high intelligence do not intend to go to college, whereas many of low intelligence have such ambitions.
10. Intelligence tests may be used in a disciplinary sense, to motivate school work, to determine the student load, to help in determining dismissal for low scholarship.

11. Surveys by means of intelligence tests show very striking differences between schools.
12. Intelligence tests are necessary in order justly to evaluate pupil progress and teaching efficiency.
13. In much educational research intelligence tests are now considered necessary, especially for the purpose of equating groups.
14. The army testing showed marked differences in the intelligence ratings of men in different occupations.
15. Intelligence tests are of use in vocational selection, but they alone cannot determine the right man for the right job.
16. In general the mentality of the delinquent is below that of the normal child.
17. Lack of intelligence is one of the causal factors at the basis of delinquent conduct, but is by no means the sole cause of such conduct.
18. Deaf and blind children in general fall below our norms in intelligence tests for ordinary children, even when compared on tests making allowance for their special handicaps.
19. Racial or national comparisons of intelligence ratings are difficult to make, and no adequate ones of unselected samplings in their respective countries have so far appeared. Our best results are based on racial groups in this country.
20. In general the comparison between American whites and American negroes shows that about 25 per cent of the negroes reach or exceed the median score of the whites.
21. Italian children in this country fall markedly below the norms on verbal tests, but not so much so on non-language tests.
22. Of other racial groups in this country we have less reliable measures of their abilities.

Review

True-False Statements

As a review mark the following statements true or false:

1. After pupils have been classified into grades, they should be divided into sections according to mental age.
2. A pupil with an I.Q. of 80 is not likely to succeed in high school algebra.
3. Intelligence is the only factor making for success in school.
4. The higher the social status of parents in general the higher the intelligence of the children.
5. A man of 35 years having a mental age of 16 on the Stanford Revision is below normal in intelligence.
6. The average score of unselected American negroes is lower than the average score of unselected American whites on intelligence tests.
7. Intelligence tests are useful in school because they measure the degree of success the teacher has achieved in getting her subject across to her pupils.
8. The homogeneous grouping of children according to intelligence widens the educational opportunity for each child.
9. Even a child with an I.Q. of 120 or above may have a special disability in a certain school subject.
10. Between intelligence test scores and general school success, there is a substantial correlation.
11. A physical handicap, such as deafness or blindness, is usually counterbalanced by keener intelligence.
12. If a child is found to have a mental age above that of the grade he is in, he should be promoted at once.
13. In comparing the intelligence of pupils in several different schools, the same tests should be used for all schools.
14. The I.Q. is adequate for accurate vocational guidance.
15. In a regrading program the M.A. is a more important consideration than is the I.Q.
16. Children with I.Q.'s of 80 and below should be encouraged to take up practical vocational training.
17. Promotions should always be made if tests show a higher degree of intelligence than the grade indicates.

18. So far as capacity is concerned most persons can pursue only one occupation with a good chance of success.

Advised Readings

Irwin, E. A., and Marks, L. A.: *Fitting the Child to the School*, New York, 1924.

Dickson, V. E.: *Mental Tests and the Classroom Teacher*, Yonkers, 1923.

Pintner, R.: *Intelligence Testing*, New York, 1923. Chapters 8, 9, 11, 12.

Poffenberger, A. T.: *Applied Psychology*, New York, 1927.

References

1. Book, W. F.: *The Intelligence of High School Seniors*, New York, 1927.
2. Burt, C.: *The Young Delinquent*, New York, 1925, Chapter 14.
3. Healy, W.: *The Practical Value of Scientific Study of Juvenile Delinquents*, Children's Bureau, Publication No. 96, Washington, 1922, page 19.
4. Irwin, E. A., and Marks, L. A.: *Fitting the Child to the School*, New York, 1924, page 134.
5. Pintner, R.: "Psychological Survey of Schools for the Deaf," *American Annals of the Deaf*, Volume 72, No. 5, pages 377–414.
6. Pintner, R.: "Non-Language Tests in Foreign Countries," *School and Society*, Volume 26, 1927, pages 374–376.
7. Pintner, R., and Noble, H.: "The Classification of School Children According to Mental Age," *Journal of Educational Research*, 1920, Volume 2, pages 713–728.
8. Scott, W. D., and Clothier, R. C.: *Personnel Management*, Chicago, 1923, Chapter 14.
9. Terman, L. M., et al.: *Genetic Studies of Genius*, Volume 1, Stanford University Press, 1925.

CHAPTER VII

THE INHERITANCE OF MENTAL TRAITS

The General Problem of Inheritance. — That physical traits such as height, color of eyes or hair, or skin, general build and similar characteristics are inherited is taken for granted by all of us. That they are inherited has been demonstrated repeatedly. That children should not inherit the physical characteristics of their parents would be considered strange. When they do not, as in the case of a Cretin or Mongolian Idiot, some special cause or reason is sought for as an explanation.

To some extent, likewise, the inheritance of mental traits has been assumed. A biographer may attempt to find in the ancestry of his subject the clue to certain character qualities and personality traits. This is frequently the case with very marked temperamental qualities or extreme character traits. The same thing is more or less true in the common experience of daily life. It is particularly true when we are at a loss to find a ready explanation of these character qualities in terms of environment. Where we cannot see an obvious causal connection between a certain type of behavior and the environment of the child, we are apt to throw upon heredity the burden of the cause.

Nevertheless there is no such general and widespread belief in the inheritance of mental characteristics as in the case of physical ones. As we have suggested, we frequently seek the cause first in the environment. With regard to some mental traits, such as general intelligence, there is frequently much doubt expressed as to its inheritance in the same way as physical traits. Since knowledge and skill are acquired in school, many people are likely to suppose that the forces of education are sufficient to account wholly for the amount of knowledge and skill acquired by the child. However, the tremendous differences in attainment revealed by children subjected to the same kind of schooling make us doubt the supreme potency of environment. We, therefore, shift to the home and try to explain the differences by means of differences in opportunities in home environment. But again we are confronted by tremendous differences in achievement among children brought up in seemingly similar home environments.

We may consider our problem as centering around two questions: (1) Are brothers and sisters of the same family (siblings) more like each other than children taken at random? and (2) Are siblings even if brought up in different environments, more alike than unselected unrelated individuals? Answers to either or both of these questions will help us to understand better the potency of heredity in regard to mental traits.

Pioneer Work of Galton. — The pioneer work in this field was done by Galton in the nineteenth century. He mapped out the two general trends that have been followed since then. In some of his studies Galton was interested in the characteristics of certain groups of related families. If in the study of such a family history a certain trait keeps appearing in one generation after another, it may be that such a trait is inherited. The chances that it is inherited become greater the more persistent it is in a family group, and the less obviously it can be accounted for by the environment in which the family has lived. In another series of investigations, Galton started with a certain number of eminent and a similar number of average men and counted their eminent relatives. The eminent group had 535 eminent relatives and the average group only 4. He also compared parents and children in certain physical and mental traits, and found as great a resemblance in the mental as in the physical traits. In this way, therefore, Galton suggested two lines of study: (1) Measurements of Resemblance by means of Correlations; (2) Family History Studies.

Correlational Studies. — Correlational studies attempt to show the amount of resemblance of any trait existing between pairs of related individuals as contrasted with the amount of such resemblance in random pairs of unrelated individuals. Such random pairs of unrelated individuals will show no correlation, just because by chance any amount of one trait may be paired with any amount of another trait. If, however, there are factors at work, such as heredity or similar home training, that make for resemblance between siblings, then the correlation between siblings should be greater than zero. Such positive correlations as we have suggested do not prove the heredity of the traits in question. If, however, the correlations for mental traits resemble the correlations for physical traits, not subject to environmental influences, then the suggestion is strong that heredity is a causal factor in the mental as in the physical traits. If, furthermore, we can find a number of cases of siblings reared apart showing correlations as high as those of siblings reared together, then our belief in the inheritance of such traits is strengthened. And, negatively, a lack of correlation among unrelated children reared in the same environment would seem to indicate the potency of heredity as contrasted with environment.

There are many studies of the resemblance of traits and we cannot here review them all. We shall merely try to indicate the general results. These results are expressed in terms of correlation coefficients and we must remember that the nearer the coefficient approaches 1.00 the greater the resemblance and that zero or very low coefficients indicate lack of resemblance. We must remember also that the amount of resemblance will not be perfect because children in the same family (or siblings) do differ from one another. And finally for the purposes of comparison we may note that correlations of from .44 to .62 have been found for such physical traits as color of eyes, color of hair, cephalic index, height, deafness and the like—traits which are hardly, if at all, dependent upon environmental influence.

Character Traits. — In this field we have little evidence, because of the difficulty of securing reliable quantitative measures. The ratings of children by teachers have been resorted to in lieu of objective measures. Teachers rated about 2000 pairs of siblings on such traits as vivacity, self-assertiveness, introspection, popularity and the like. The average of 18 correlations is .53. If the accomplishment quotient (that is the E.Q. divided by the I.Q.) can be taken as an indirect measure of effort put forth by the child, we find a correlation of .32 for 188 pairs of sibs tested on intelligence and achievement tests.

Now that objective tests of character traits are becoming available, more reliable measures are beginning to appear. Investigations (May and Hartshorne, 5) of sibling resemblance by means of objective tests for the measurement of deception show a correlation of .47. It would seem, therefore, that the resemblance of siblings in character traits is about as large as that found for physical and for intellectual traits.

Intelligence. — Here we have a great many more studies, and all of them made by means of objective tests. The average of several correlations reported by early workers in this field who used psychological tests such as cancellation, memory and the like is about .30. The average of twelve later reports using standard intelligence tests is about .51, based upon about 5000 pairs of siblings, with a range of correlations from about .33 to .68. It would seem as if the correlation increases with increase in accuracy and adequacy of the trait measured.

School Achievement. — There are also many studies of the resemblance of siblings in school work. The measures in these cases are teachers' grades, general scholastic records or special tests of specific studies. The average of the many correlations reported is again .50.

Parent-Child Relationship. — Two studies report intelligence tests given to parents and children, giving average correlations of .35 and .53. These correlations are much the same as those found between siblings.

Twins. — The resemblance of twins in intelligence has been studied by several workers and if we sum up the correlations obtained we find that the average is about .78 for about 400 pairs. The correlations here are decidedly higher than for siblings. There seems also to be some evidence that so-called identical twins resemble each other on intelligence tests somewhat more closely than do non-identical twins.

There are several detailed studies of the specific reactions of twins on a given series of tests. Figure 37 shows the profile of the reactions to the Stanford Binet Test by one pair of twins. The failures and passes on the test occur at almost identical places throughout.

Cousins. — Studies of further degrees of relationship are very scarce. A few correlations have been reported for cousins with reference to school achievement and intelligence. The average correlation for about 500 pairs is .23.

If now we look at these three degrees of relationship, twins, siblings, and cousins, we note that the correlations decrease as the relationship becomes less close: about .70 for twins, about .50 for sibs, about .20 for cousins and zero for unrelated children. There is great likelihood that heredity is a potent factor in causing these resemblances. Furthermore, the resemblance for siblings in intelligence,

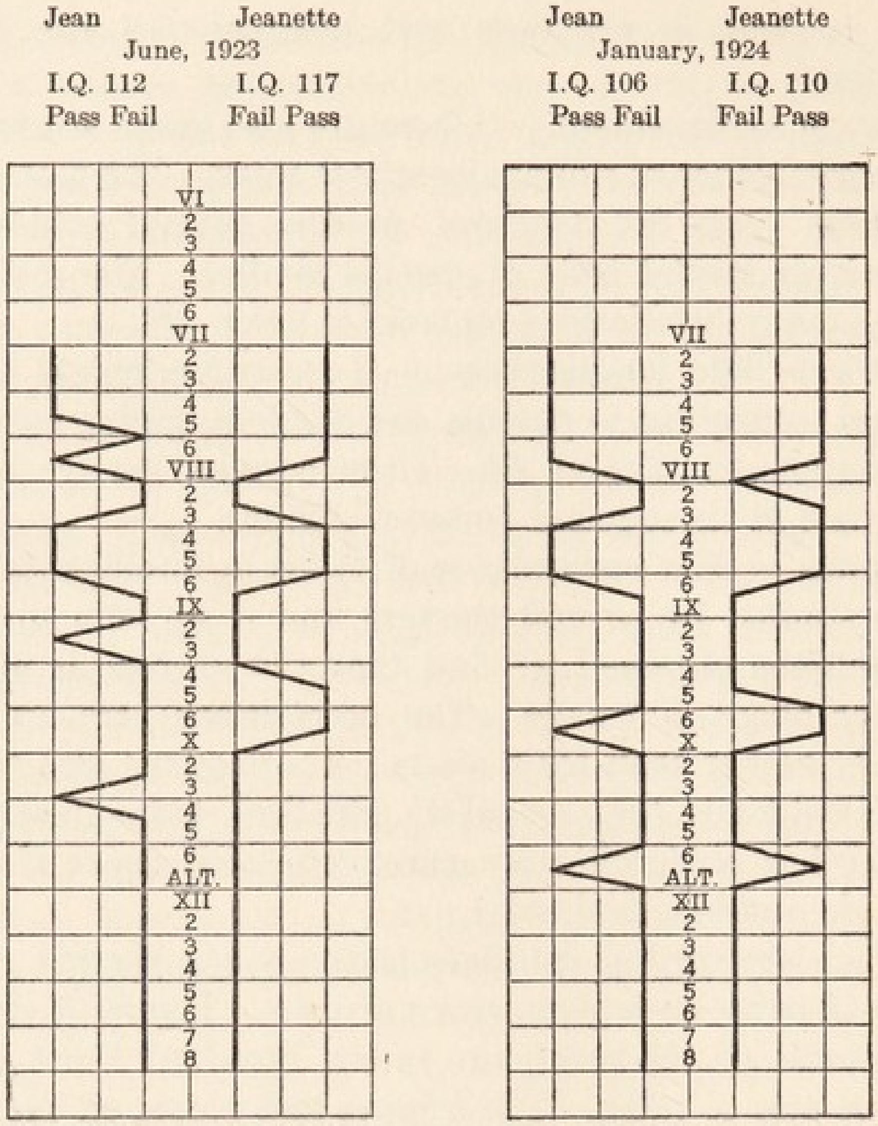


FIGURE 37. — Stanford-Binet examinations of twins. (From Wentworth's *Individual Differences in the Intelligence of School Children.*)

character and school achievement is about the same as the resemblance in physical traits. Since the resemblance in physical traits cannot be considered as due to environment, it is argued that the resemblance in mental traits, as in physical traits, is primarily a consequence of heredity.

Dependent and Foster Children. — Nevertheless, these correlations do not prove that heredity is the sole factor. A similar environment might be influential in causing these correlations, for we note that as the correlations increase from unrelated children up through cousins and sibs to twins, so also does the similarity of environment for each pair increase. These correlational studies cannot tell us how much of this resemblance is due to heredity and how much to environment. Other studies (5) have, therefore, been made to find out what influence on intelligence and achievement ratings has been produced by changes in the environment of children. These studies deal with the changes produced in dependent children adopted into foster homes, with comparisons between siblings reared apart in different homes and siblings reared together, and with unrelated children reared together. These studies are recent and somewhat contradictory. It is difficult to evaluate them at the present time, but in general they show that some increase in I.Q. on such a test as the Binet takes place in young children when transferred from a poor home to a better home. We must not, therefore, think of a Binet I.Q. as a fixed unchangeable measure of a child's intelligence. By a change to a superior environment this rating can be increased from 10 to 20 points, if the change takes place at an early age and if it is from a very poor to a very superior environment. Such changes, however, are rare. They are more likely to take place if the child is moved to the good environment when very young, and less likely to take place with children of school age.

Family History Studies. — The second method of showing the inheritance of mental traits is to analyze the characteristics of a family group and to note what mental traits predominate. If there is a great similarity in certain traits through several generations the chances are that this is due to the inheritance of underlying basic abilities or tendencies to reaction. Galton (2) stimulated such studies by his early work and his own family itself illustrates well the concentration of eminent individuals within one family group. He himself was the cousin of Charles Darwin and the grandson of Erasmus Darwin, a distinguished botanist and author. Of the two sons of Erasmus, one was distinguished in experimental research and the other was a notable physician and father of Charles Darwin, one of the greatest men in modern science. He married Emma Wedgwood, a granddaughter of Josiah Wedgewood of pottery fame. Charles Darwin's four sons all became notable, one as a botanist, another as an astronomer, another as an engineer and another as a political economist and eugenist. Figure 38 is a family history chart showing the eminent people in these related family groups. This figure illustrates the typical method followed in family history investigations.

Another famous family whose descendants have been traced is the Edwards family. Johnathan Edwards was born in 1703 and in 1900 there were 1394 descendants accounted for. Among these we find 12 college presidents, 60 physicians, 65 college professors, about 100 clergymen and missionaries, 75 army or navy officers, 60 prominent writers, by whom 135 books of merit were written and 18 important periodicals edited, about 100 lawyers, 30 judges, 80 public officials, such as United States Senators, governors, members of Congress, mayors, and ambassadors, as well as many successful bankers, business men and managers of large industrial enterprises. There is no trace of any feeblemindedness in this group and no record of any criminal.

On the other hand there are also records of families in which the opposite type of individual predominates. One of the earliest of such studies is that of the so-called Jukes family (1). Max Juke was born in 1720. He was a shiftless, intemperate and illiterate backwoodsman, married to a stupid worthless woman. Up to 1877 about 1200 descendants had been traced, and among these we find 310 paupers, 7 murderers, 60 thieves, 50 prostitutes and 130 convicted of crime. Only 20 of the family learned a

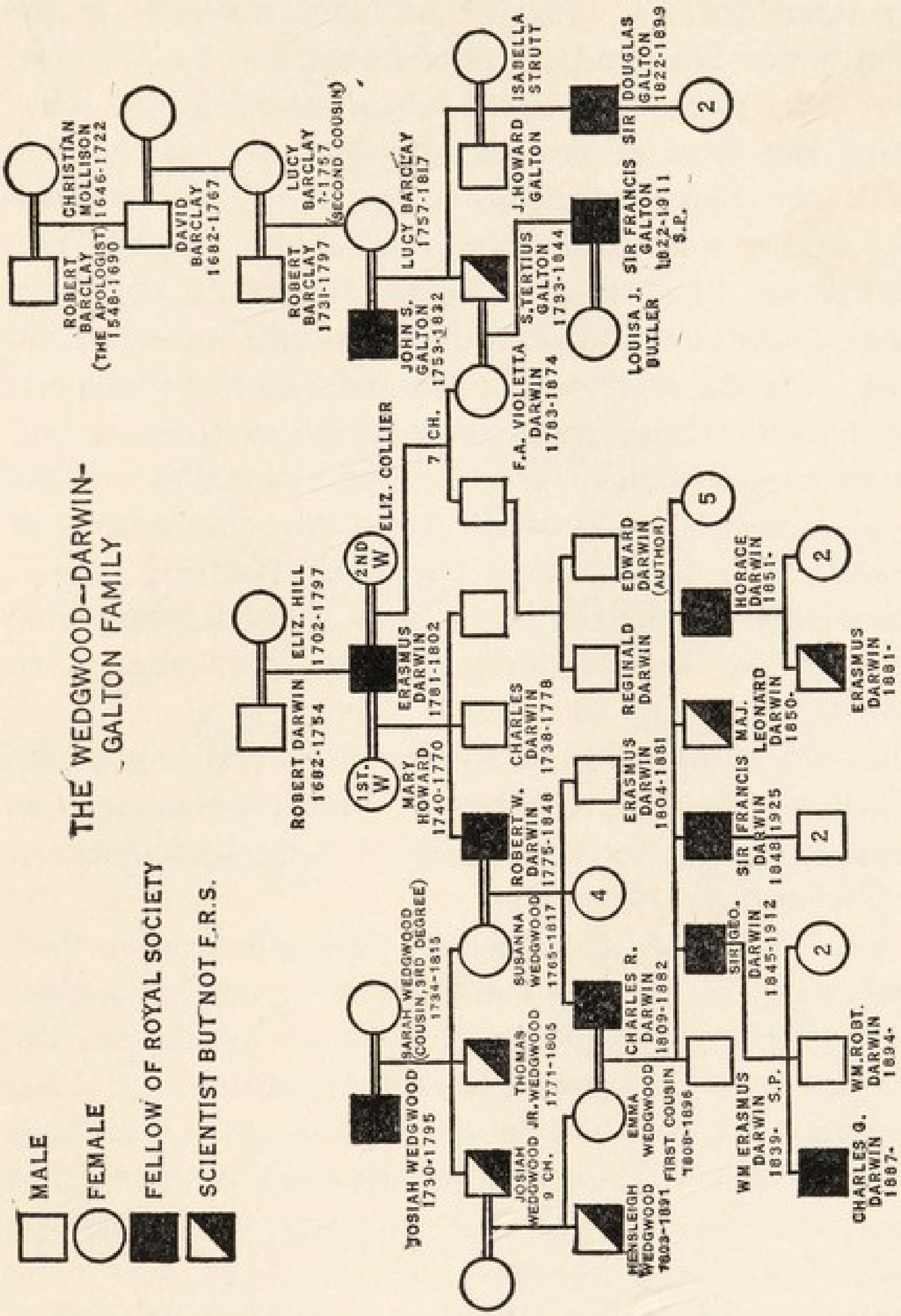


FIGURE 38. — A family history chart showing many eminent people. (From Sandiford's *Educational Psychology.*)

trade and 10 of these did so in prison. A later follow-up in 1915 of the Jukes showed that the descendants still manifested the same traits and that migration of some branches of the family to a different environment had not changed their original native tendencies.

Another striking picture is that offered by the Kallikak family (Goddard, 3). Martin Kallikak, a soldier in the Revolutionary War had an illegitimate son by a feebleminded girl. Later he married into a good family. Thus we have two lines of descendants from the same man. In 1912 the illegitimate line showed 480 descendants among whom were 143 feebleminded, 36 illegitimates, 33 prostitutes, 24 alcoholics, 3 epileptics and 3 criminals. The legitimate line showed 496 descendants, mostly good citizens including doctors, lawyers, educators, judges and business men. Only five were reported as showing bad traits, such as alcoholism, insanity and immorality. There were no criminals. Figure 39 shows the descendants of the legitimate and illegitimate lines of the Kallikak family. In this picture we note the great number of feebleminded individuals and the tendency for the feebleminded to marry the feebleminded.

There are many other family histories pointing in the same general direction as those we have mentioned. The tendency of certain traits to persist in a family is a strong presumption of heredity.

In spite of all the changes that take place in families, we seem to see no marked change in their personal characteristics—the intelligent family remains intelligent, the dull family dull. If opportunities came in the way of the dull family, and probably some did, they were unable to make use of them. If misfortunes overtook the intelligent families, and probably some did, they were able to survive and overcome them.

Relatives of Superior Children. — Similar but somewhat different evidence of the inheritance of desirable traits is shown in Terman's (7) study of the relatives of superior children. Of the 62 members of the Hall of Fame, 14 or 22.6 per cent are known to be related to one or more of the superior children studied. Twelve of the 643 children have parents or grandparents in *Who's Who*, even although the average age of the fathers is only 41

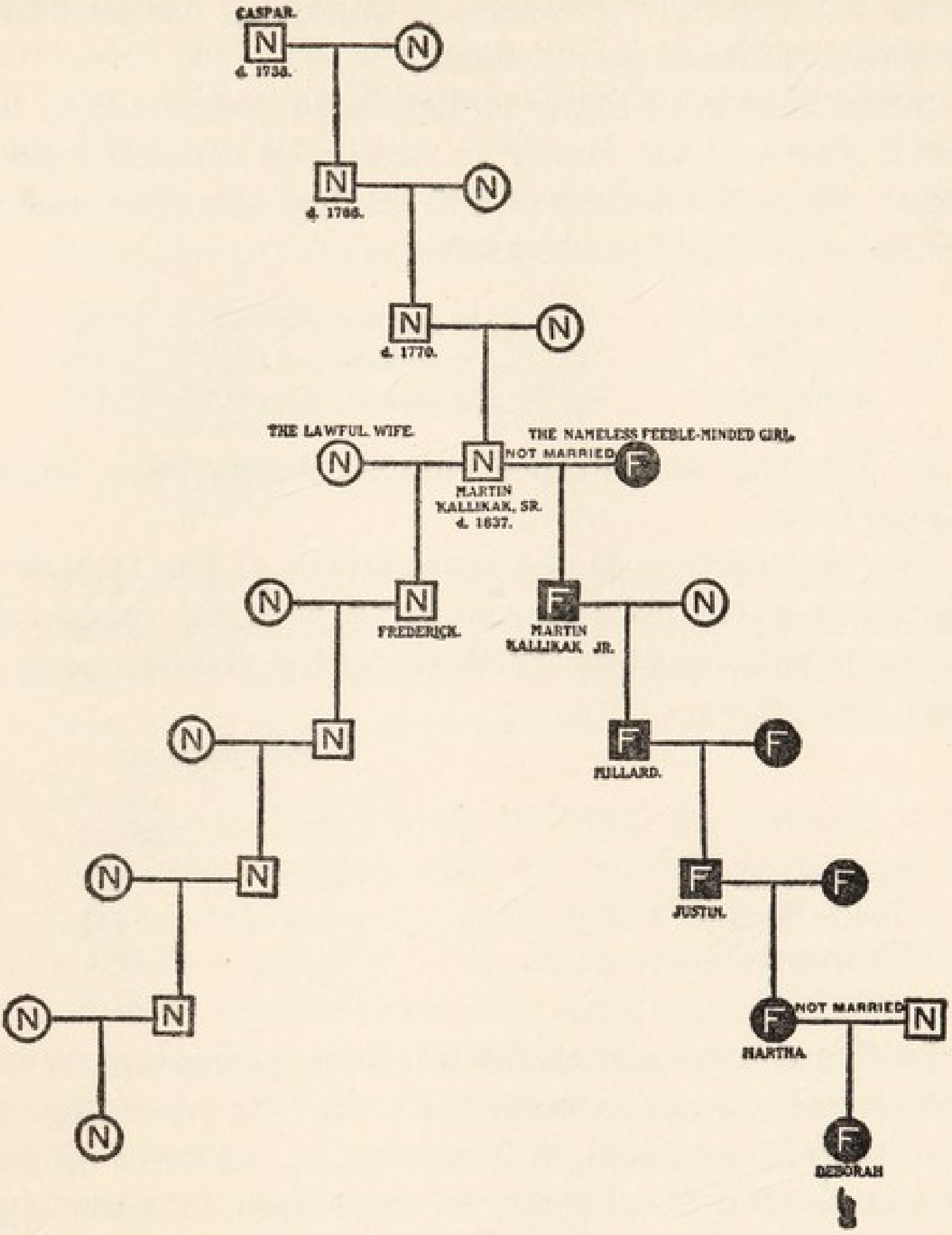


FIGURE 39. — The legitimate and illegitimate lines of the Kallikak family. The black denotes feeblemindedness. (From Goddard's *Kallikak Family.*)

years, whereas the average age of first inclusion in *Who's Who* is 49.9 years. There were 35 other relatives reported in *Who's Who.* In addition there were 58 other relatives of sufficient distinction to be named in standard cyclopedias of biography. Parents and grandparents have held many responsible positions, 46 holding important national or state positions, 67 important religious or fraternal positions, 27 university positions, 74 important business positions and 18 bank presidencies.

Social Status. — Many studies have been made of the social status of the family in which the eminent man is born. Thus the fathers of 885 leading American men of science were distributed as follows:

Professional	43.1%; in general population 3.0%
Business	35.7%; in general population 34.1%
Agriculture	21.2%; in general population 41.1%

In England and France similar studies have shown similar facts.

The distribution of the occupations of the fathers of Terman's gifted children compared with the proportion in the population from which the gifted children came is very instructive.

	Percentages	
	Fathers of Gifted	*General Population*
Professional	29.1	2.9
Public Service	4.5	3.3
Commercial	46.2	36.1
Industrial	20.2	57.7

In this comparison notice the large percentage of the professional group as contrasted with its percentage in the general population, and conversely, the small proportion in the industrial group as contrasted with the large proportion in the general population. Eminent men not only tend to come from the professional classes, but these professional classes show a much greater proportion of gifted children. What makes for eminence in after life evidently tends to show in early childhood.

Another way of looking at this problem is shown by a comparison of the intelligence ratings of different occupational groups. One study by Sandiford (6) includes about 5000 high school and college students. Another study by Haggerty and Nash (4) includes about 8000 children of all grades. The tests used were not the same, so we need pay little attention to the absolute values of the I.Q's. Below we have combined the results of these two studies:

	Median I.Q. of Children	
Occupation of Fathers	*High School*	*All Grades*
Professional	105	116
Business	103	107
Skilled	103	98
Semi-skilled	102	95
Farmers	102	91
Unskilled	101	89

The important point to notice is that the medians, although different, decrease in the same fashion as we descend from the professional to the unskilled group. There are many other studies of this nature all showing the same general trend.

Of course this relationship between the occupation of the parent and the intelligence of the child is no proof of heredity, but taken along with the other facts of this chapter it fits in admirably. The more intelligent children tend on the whole to have the more intelligent parents and these again tend to be found in greater numbers in the higher occupations. For the higher the occupation, the more likely is it to require intelligence. The medians or averages quoted merely show the general tendency of a rise in intelligence from the unskilled to the professional group. If the whole distributions were to be shown, we would see that there is much overlapping from group to group. There are some in the unskilled group with intelligence as good as, or better than, the average of the professional group and, conversely, some in the professional group with no better intelligence than the average of the unskilled.

This general hierarchy of occupations with reference to the intelligence of the men in them can be seen in Figure

32 on page 150 of this book. It is the chart of the intelligence of occupational groups in the army, which we used in the last chapter in another connection. The diagram shows the median intelligence rating by a short vertical line, and the length of the line shows the distribution of intelligence for the middle fifty per cent. We must remember, therefore, that twenty-five per cent of the cases lie to the right and twenty-five per cent to the left of the horizontal line. The letters, A, B, C+, etc., indicate the intelligence ratings. There were a few laborers who achieved an A rating. The middle fifty per cent, however, ranged from C to D and the lowest twenty-five per cent fell in the D and D− classes. At the upper end notice that the middle fifty per cent of the medical officers lie in the B and A classes, that the upper twenty-five per cent is wholly in the A class, while the lower twenty-five per cent lies in the C+ or lower classes. As a matter of fact less than one per cent fell below C and there were none in the D and D− classes. The diagram then illustrates well the general tendency to increase in intelligence as we go from laborer to professional man, and it also illustrates the wide range of intelligence to be found at each step. Much more than intelligence determines a man's occupation, but intelligence is one of the factors. If, then, intelligence is inherited we should expect to find a similar hierarchy in the sons of the men in these different occupations and such a general hierarchy, as we have already shown, exists among school children.

Summary

1. Mental characteristics would seem to be inherited in the same manner as physical characteristics.
2. The pioneer work in the study of the inheritance of mental traits was done by Galton.
3. He started both methods of investigation, namely (1) correlational study of resemblance and (2) the family history study.

4. Correlations of siblings for physical traits range from about .44 to .62.
5. The average of many correlations for character ratings of siblings by teachers is .53.
6. The average of many correlations for intelligence tests of siblings is .51.
7. The average of many correlations for scholastic tests and grades of siblings is .50.
8. The correlations for the mental traits of siblings are about the same as the correlations for the physical traits.
9. There are very few studies of the resemblance of parent and child. Correlations of .35 and .53 have been reported.
10. The correlations of intelligence tests of twins average about .78.
11. A suggestion of the resemblance between cousins is indicated by one report of a correlation of .23.
12. Random samplings of unrelated children show zero correlations on intelligence tests.
13. The correlation would seem to decrease the further apart the degree of relationship becomes.
14. A marked change in environment begun at an early age and continued for several years may lead to a change in intelligence as rated by the Binet Scale.
15. Family history studies show desirable or undesirable traits running through many generations.
16. The relatives of children testing high on intelligence tests are on the whole very decidedly superior to the relatives of the average run of people.
17. There is a decided correlation between the occupation of the father and the intelligence rating of the offspring.

Review

True-False Statements

As a review mark the following statements true or false:

1. Galton found that a group of eminent men had no larger number of eminent relatives than a group of average men.
2. In mental tests, the resemblance between twins is no greater than that between siblings.
3. Ancestry reduces the variability of offspring and determines the point about which they vary.
4. There is little correlation between the general intelligence of children and the social status of their parents.
5. The resemblance between siblings in I.Q. is greater than that between unselected, unrelated individuals.
6. Correlations between siblings in mental traits are in general higher than the correlations in physical traits.
7. The resemblance between siblings in school achievement is much greater than in intelligence tests.
8. Foster children seem to improve somewhat in I.Q. if adopted early and brought up in a very good environment.
9. The Kallikak study of Goddard proves that feeblemindedness is hereditary.
10. It was the poor environment surrounding the Jukes that led to so much poverty, criminality and immorality.
11. If the Jukes had been transplanted to a better environment, the chances are they would have proved perfectly normal.
12. Every child with an I.Q. above 140 is almost certain to have eminent relatives.
13. Eminent men tend to come from the professional class, and this class produces a relatively greater number of children with high I.Q.'s.
14. Underline the best responses:
    The correlation between siblings in intellectual traits is about 10, 20, 50, 75, 90.
15. The correlation between twins in intellectual traits is about 10, 20, 50, 75, 90.
16. The highest average I.Q. is likely to be found among children whose fathers are: (1) laborers (2) farmers (3) professional men (4) business men.

Advised Readings

Galton, F.: *Hereditary Genius*, New York, 1869.

Goddard, H. H.: *The Kallikak Family*, New York, 1914.

East, E. M.: *Heredity and Human Affairs*, New York, 1927.

References

1. Dugdale, R. L.: *The Jukes*, New York, Fourth Edition, 1910.
2. Galton, F.: *Hereditary Genius*, New York, 1869.
3. Goddard, H. H.: *The Kallikak Family*, New York, 1914.
4. Haggerty, M. E. and Nash, H. B.: "Mental Capacity of Children and Parental Occupations," *Journal of Educational Psychology*, 1924, 15: 559–572.
5. National Society for the Study of Education, Twenty-Seventh Yearbook, 1928.
6. Sandiford, P.: "Paternal Occupations and Intelligence of Offspring," *School and Society*, 1926, 23: 117–119.
7. Terman, L. M. et al.: *Genetic Studies of Genius*, Volume I, Stanford University Press, 1925, Chapter V.

PART II — THE MODIFICATION OF ORIGINAL NATURE

CHAPTER VIII

THE LAWS OF LEARNING

General Tendency of Modification. — So far, in the first part of this book, we have been trying to find out what the essential nature of the individual is, what his original nature is. In doing this we have made a survey of his capacities and tendencies, and we have laid especial emphasis upon any instruments that have been constructed which attempt to measure these capacities and tendencies, because such attempts at measurement give us a more exact knowledge of man's original nature. Now there is one aspect or tendency of man that has not been specifically mentioned, but which has been taken for granted in all our previous chapters, and this is his tendency to change or be modified within limits by his surroundings. It is with this aspect of man's original nature that all the rest of the book will deal. It is at once the most common-place fact about human nature and at the same time the most important, surprising and significant fact. If this tendency of man's conduct to be modified or changed did not exist, there would be no education and no such study as educational psychology. By original nature man is a modifiable organism, his inherited tendencies set the bounds within which this modification shall take place and his immediate environment determines the type of modification that will be made. If a child is of normal intelligence he will learn to speak English if born in England, Spanish if born in Spain, and Chinese if born in China; but if he inherits the mentality of an idiot, he will learn no language, and the three different language environments we have mentioned would not cause this change or modification in his behavior. Some kinds of environment will affect or modify some individuals, while other kinds of environment are needed to modify other individuals. There is thus a continual interplay between the individual and his environment. What type of environment shall be effective will be determined by the original nature of the individual, and what specific modifications are to be made, granted an effective environment, will be determined by the environment. In what follows we shall assume the limitations imposed by original nature and emphasize merely the effect of the environment in causing changes.

Life is a never-ending series of situations confronting the individual, a continuous series of stimuli. To this never-ending stream the individual is continually reacting and each reaction that the individual makes leaves its impression upon the individual. It changes or modifies him even if ever so slightly. An "experienced" man is a man who has had lots of "experiences," a man who has reacted to many different stimuli. If you are experienced in a certain field, you have made a lot of different reactions, and you have been changed by those reactions, otherwise you would be just as green as the veriest "greenhorn" in that particular field.

Necessity for Reaction. — Life is reaction. If there were no reactions, the individual would be dead. A continual series of reactions is necessary in order to maintain one's equilibrium in the face of the ever-shifting environment. When the situation changes suddenly the organism is called upon to make a number of sudden reactions to maintain its equilibrium. This is obviously true physically, as, when one slips with his foot, a number of quick physical adjustments may be necessary to maintain one's upright posture. It is just as true mentally for a situation may arouse painful feelings or ideas and we react vigorously in all directions in order to remove the discomfort and regain our equilibrium.

We keep reacting, therefore, in order to maintain a certain equilibrium or satisfying state of affairs. So long as a situation is annoying to us we are spurred on to repeated reactions in order to remove the annoyance and return to our former state of satisfaction. Satisfaction and annoyance, therefore, are the great spurs to action. They are the supreme teachers. Unless we can remove the annoyance and attain a state satisfying to us, we have not learned our lesson. When we attain a satisfying state, we have learned how to meet a given situation and we try to repeat this reaction again on the recurrence of the same situation.

Tendency to Seek Satisfaction. — This tendency to return to a certain equilibrium (Raup, 1) is common not only to animals and plants, but to the inorganic world as well. The equilibrium satisfying to any particular organism will be determined by the inherited constitution of that organism. In this way, then, we have the mechanism set up for the ceaseless reactions of man and at the same time the explanation why he will react again and again to the same situation, varying his responses from one to another until he finds one that allows a complete or partial achievement of a satisfying equilibrium or until the situation shifts again and he is called upon to react in another way. It is in this repetition of reactions, this multiple response to the same situation that opportunities for modification occur. We are thrown into a state of annoyance, we try to adjust, we are partially successful and then the situation changes. We think over what has happened, trying out other possibilities in our imagination. The situation occurs again, we try these other possibilities and are more successful and so on. Thus we have a gradual learning, a trial and error in actual behavior, a trial and error in the realm of ideas.

Trial and Error. — The occurrence of a situation, a reaction to it followed by feelings of annoyance, another slightly different reaction and so on until success or partial success accompanied by satisfaction results, this in general is the bare outline of the process of learning. Thus in learning a complex physical activity such as swimming, when confronted with the situation of standing in the swimming pool, swimming instructor giving directions and so on, the individual makes certain bodily movements, feels himself sinking and this causes great annoyance. He reacts vigorously by somewhat different movements, if these are partially successful he will tend to repeat them or similar ones the next time he is confronted with the "sinking" situation. If the particular reactions are unsuccessful and he continues to sink, he tends not to repeat those reactions when again confronted with the "sinking" situation. Thus gradually do certain reactions tend to become connected with certain situations. It is a gradual trial and error process until the individual learns a reaction that is satisfying to him.

This trial and error procedure is not only true of the learning of physical acts but is equally true of all kinds of learning. Thus in learning to recite a poem we may read it over several times and then try to repeat it from memory. We come to the end of the first stanza and try to begin the second. We say a few words and then stop. It does not sound right. We are not satisfied. We try again. Perhaps this time the words sound right and we are satisfied and so proceed, or else we are not satisfied and try some other words or give up and go back to the book in order to get the right words and start all over again.

So in playing chess we find a continuous mental trial and error. To a given situation on the chess board we react by imagining what would happen if we made a certain move. We notice that if we made this move our opponent might make either of three or four moves. We follow this on by imagining or trying out in our mind a counter move to each one of the three or four possibilities, and so on, until we are unable to envisage the changed situation. If we arrive at an imagined position of the board satisfactory to us, we make the move in question, if not we start out again on another imaginary move. We try out mentally many moves until we reach one that appears most satisfying or least annoying.

The baby says, "ta, ma, da," and the mother shouts in glee "mamma" and laughs and cuddles it, and the baby has decided feelings of satisfaction. It may repeat the sounds, "ta, ma, da" or "ta, ma, ma" and so on, because they are a satisfactory response to a certain situation. As time goes on more satisfaction (attention of the mother) results when "ta, ma, ma" is said than when "ta, ma, da" is said, and so "ta, ma, ma" becomes selected as preferable in a given situation and gradually "mamma" is found to be the most satisfying response to certain types of situations. The baby will quickly learn to say those things which lead to satisfyingness. If parents retain their great enthusiasm for the "cute" baby talk of the early stages, then this baby talk may last for years, and occasionally we meet a six or seven-year-old who still retains many baby words and baby expressions, because those around him have caused satisfaction to be felt by him in such responses and they are therefore retained.

In all improvement due to such trial and error reaction, the individual stops at a level which is satisfying to him. The mere repetition of a response will not cause improvement in that response. Improvement will only come if the result is annoying to the individual. As long as the response is satisfying to the individual, there is no motive to change the response. Thus the baby will go on talking baby talk, so long as such talk is approved by those around him. He does not imitate the speech of those around him unless he is taught to do so — taught by the universal laws of satisfaction and annoyance.

Mere repetition does not lead to improvement. This is well illustrated by the old story of Johnny who was kept in after school by his teacher to improve his English usage. Johnny had written "I have went" in a little composition and so his task was to write out "I have gone, I have gone" one hundred times. One hundred repetitions would surely stamp in the right reaction. So Johnny wrote and when he had finished the hundredth repetition, he looked up and found that the teacher had left the room. Feeling that he had done his task and had a right to go, he went up to the teacher's desk and laid his work on it. Then thinking that it would be more polite to explain his going, he wrote a little note, "Dear Teacher, I have done my lesson and seeing you were not here, I have went."

The Laws of Learning. — We are now ready to summarize our discussion into the two laws of learning: —

The Law of Exercise. — Whenever a modifiable connection between situation and response is exercised, the strength of that connection is increased.

The Law of Effect. — Modifiable connections between situation and response, causing satisfaction, are strengthened, causing annoyance, are weakened.

The first law is sometimes called the Law of Use and Disuse. The fact of disuse is, of course, implied in the statement above of the Law of Exercise. Exercise of a function strengthens it, disuse or lack of exercise weakens it. The Law of Exercise (Use and Disuse) is, however, impotent without the motive force of the Law of Effect. Reactions which are accompanied or followed by satisfying states of affairs are learned quickly or tend to be repeated. Reactions which are accompanied or followed by annoying states of affairs tend not to be learned and tend not to be repeated. We are satisfied when we have achieved our purpose or aim, we are annoyed when we fail to do so. To reach for food on the table, seize it and eat it is satisfying; to reach for food and be unable to seize it is annoying and will lead to another and modified reaction on the part of the individual.

Annoyance resulting from a given response may be so great that the individual is baffled and enraged. He feels that it is impossible to proceed and so he turns away from the given situation. He gives up. He refuses to react. This is what happens when things are too difficult for us. We are trying to solve a problem but reaction after reaction takes place with no solution so that eventually we acknowledge defeat and give up. We turn away to make other easier reactions that are satisfying to us. In the process of thinking, success is the great stimulant to the next step in the procedure. If stage one is successful we go on gleefully to stage two and so no. "Nothing succeeds like success." The glow of satisfaction in achieving one step on the way to our final goal spurs us on to take the next step. Problems set before the child must not be too difficult. He must be able to gain successes along the road to spur him on to continue to the end.

Readiness. — We have stressed the fact that a reaction must be satisfying to the individual, and it is satisfying only when his purpose has been achieved. At one time a given situation will provoke a certain response and at another time a different response and both may be satisfying to the same individual. What is one's purpose or aim at any given moment, what will cause satisfaction and what annoyance, depends upon the state of readiness of the individual to react. That an individual has an aim or purpose, means simply that he is ready to react in a given direction, that there are certain kinds of stimuli or parts of stimuli to which he will react more readily than to others, and with which he will be more satisfied. This is sometimes called the "Law of Readiness." It emphasizes the condition or state of the reacting individual. The individual is not the same from moment to moment. The total bodily condition changes and the state of the organism at any given moment is a state of readiness for certain stimuli and unreadiness for others. To respond to the stimuli for which the organism is ready is satisfying, not to respond is annoying. Suspense or expectation when long delayed is annoying, for the organism is ready to respond to certain kinds of stimuli which do not appear. The state of readiness at any given moment is determined by the whole bodily condition of the organism plus the reactions which have immediately gone before. To be forced to change suddenly from one thing to another is generally annoying because the organism is not ready for the sudden shift.

One very common cause of annoyance, leading to angry outbursts out of all proportion to the seriousness of the affair, occurs very frequently in automobile driving. The unanticipated movements of other cars cause decided annoyance and flashes of anger arise quickly. We always feel, of course, that we are in the right and that it is the other driver's fault. But a calm review of the situation reveals that frequently the anger is there although we are in the wrong. The acute feeling of annoyance or anger is undoubtedly due in many cases to the sudden necessity for changing our response. We are ready or set to continue reacting in a certain way, driving straight ahead, and when the need for sudden change arises, a feeling of annoyance springs up. Note that on a crowded street we are ready or set for frequent sudden changes, and we do not have the feelings of annoyance as frequently as on less crowded thoroughfares.

What reaction will follow the presentation of a given stimulus cannot be foretold, until we know the state of readiness of the subjects. In the school room what reaction will follow the stimulus $\begin{array}{c}3\\ \underline{2}\end{array}$? It might be 5 or 1 or 6, all depending upon the preceding work of the class. If the lesson is one in addition, the subjects will add, if subtraction then they will subtract and if multiplication they will multiply.

This readiness of the organism to react in a given direction is at the basis of those states of mind that we call motives or desires or purposes. In the field of education, it has been called the will to learn. Without this will to learn, improvement is impossible. This is the same as saying that if the organism is not ready to respond, to be made to do so is annoying and, therefore, improvement does not take place. If the child does not wish to learn his multiplication (if his organism is not ready), and we force him to repeat it twenty times, no improvement takes place. If he wishes to learn (if his organism is ready), and he repeats it twenty times, improvement will take place.

Basis of All Improvement. — The Laws of Readiness, Effect and Exercise give us the basic facts as to improvement. Readiness selects the parts of the environment to which the organism is most likely to respond. Exercise stresses the necessity for repeated reaction in order to achieve improvement. Effect selects the reactions which will survive.

Translated into terms of the school, we may say that readiness means the arousal of certain desires or aims in the pupil, exercise stresses the need for practice and effect means that satisfaction must be attached to the reactions we desire to have the pupils make. If the pupils are ready, if they want very badly to do or learn a given thing, they will make numerous reactions and they will feel great satisfaction at reaching a successful conclusion. Hence the importance of motivation in all learning. The motives may be of different kinds, desire to beat one's neighbor, to avoid blame or punishment, to please the teacher or parent, to make or construct a certain thing, to help in being the best class in the school, to help the school or community. There must be some motive or else readiness will not exist and if it does not satisfaction will not result from the desired reaction.

A teacher, therefore, must see to it that the pupils are in a state of readiness for the subject-matter to be learned, that they receive sufficient exercise in reacting to it with a clear knowledge of the results of their reactions as they are learning, and lastly, that the effect of their reactions be immediately felt by them in terms of satisfaction or annoyance. Try to get the pupils interested in their work. Keep records of their improvement. Correct the examination papers as soon as possible, and tell the pupils their marks. Praise liberally for good work. Be sparing of blame or punishment. If necessary punish right after delinquency so that the annoyance may be connected with the faulty reaction in question.

Summary

1. By original nature man is a modifiable organism; he possesses the tendency to be changed by his reactions.
2. All education rests upon this modifiability of the human organism.
3. Life is a never-ending series of situations to which the individual is called upon to react.
4. We keep reacting in order to obtain satisfaction and avoid annoyance.
5. We react again and again (trial and error) until we arrive at a state satisfying to us.
6. Improvement does not take place by repetition alone, but only by seeking satisfaction.
7. The Law of Readiness states that when the individual is ready to react in a certain direction, to do so is satisfying and not to do so is annoying.
8. The Law of Exercise states that whenever a modifiable connection between situation and response is exercised, the strength of that connection is increased.
9. The Law of Effect states that those reactions accompanied by or followed by a satisfying state of affairs tend to be repeated; and conversely, those accompanied by or followed by an annoying state of affairs tend to be inhibited.

Review

True-False Statements

As a review mark the following statements true or false:

1. Readiness is determined by all one's original tendencies and all one's experience up to date.
2. If we wish to strengthen a connection it is sufficient that we exercise it.
3. If a child is not interested in a subject, he will not be ready to respond.
4. Satisfaction is not equivalent to pleasure, nor annoyance to pain.
5. Ordinarily individuals do not differ very much in readiness to respond.
6. Reactions followed by satisfaction tend to be repeated and learned.
7. Painful stimuli are always annoying.
8. A definite purpose makes us ready to respond to certain stimuli rather than to others.
9. Much drill in school is the only guarantee for improvement.
10. Annoyance should be attached to a response if we wish to eliminate it.
11. One of the chief functions of a good teacher is to make the child ready to respond to desirable stimuli.
12. Teaching is the arrangement of situations which leads to desirable bonds and makes them satisfying.
13. Regardless of much exercise, the beauty of our handwriting does not increase because we are satisfied with it as it is.
14. Knowledge of success or failure helps the learner in his learning.
15. The "set" or attitude of the learner helps to determine which reactions will be satisfying.
16. Practice always makes perfect.
17. Complete the following sentence:
    Reactions followed by ——— tend to be ———; those followed by ——— tend not to be ———.

Advised Reading

Perrin, F. A. C. and Klein, D. B.: *Psychology*, New York, 1926, Chapter V.

Reference

1. Raup, R. B.: *Complacency*, New York, 1925.

CHAPTER IX

THE LEARNING CURVE

Making a Learning Curve. — In the previous chapter we discussed the laws governing improvement. In this chapter we shall study the way in which improvement actually takes place when we practice or repeat a reaction. The best and quickest way to understand what takes place is for the student to construct a learning curve for himself. The student is strongly recommended, therefore, to pause here in his reading and to carry out the following exercise.

Writing Alphabet Backwards. — Get several loose sheets of paper and a watch with a second hand (or better still a stop watch). You are to write the alphabet backwards 30 times without allowing yourself to make any mistake. If you write down the wrong letter, cross it out and correct it. Your score will be the time taken to finish the alphabet backwards correctly. Use small letters and do not join them one to the other.

Put 1 on the first line of your paper, denoting the first trial. Note the time of starting on the first line of your paper and immediately write the alphabet backwards beginning with z and going back to a. As soon as you have finished note the time and write the time taken in seconds at the right hand side. Then read through the alphabet forwards to be sure you have made no mistake. If you have, penalize yourself by adding five seconds to the time, but try not to make any mistakes, correcting as you go along. This is perfectly possible. Fold the paper over backwards so that you cannot see what you have written. Continue in this way for thirty trials without stopping.

Making a Graph. — Now make several graphs of your results as in Figure 40. Figure 40 is constructed from Table 2, which gives the number of seconds for each in-

TABLE 2

TIME IN SECONDS FOR WRITING THE ALPHABET BACKWARDS: THIRTY TRIALS

Trial No.	*Time*	*Trial No.*	*Time*
1	41	16	18
2	38	17	19
3	34	18	17
4	30	19	18
5	26	20	15
6	24	21	16
7	23	22	17
8	22	23	19
9	23	24	17
10	19	25	15
11	21	26	15
12	21	27	14
13	20	28	15
14	22	29	14
15	20	30	15

dividual trial. The numbers along the base line show the trials from 1 to 30. The numbers along the vertical line show the time in seconds from zero to 45. On your chart make allowance for more than 45 seconds, if your first trial took longer than 45 seconds. Notice how the curve in Figure 40 drops very quickly at the start showing considerable improvement for the first eight trials. The ninth trial is poorer than the eighth, but the tenth trial shows improvement again. From this point on the curve is very irregular. There is no consistent gain from one trial to the next. There is, however, a general tendency towards gain up to trial 27. The last three trials, 28, 29, and 30, show no improvement over trial 27. We can see

FIGURE 40. — Practice Curve. Writing the alphabet backwards. Time of each trial.

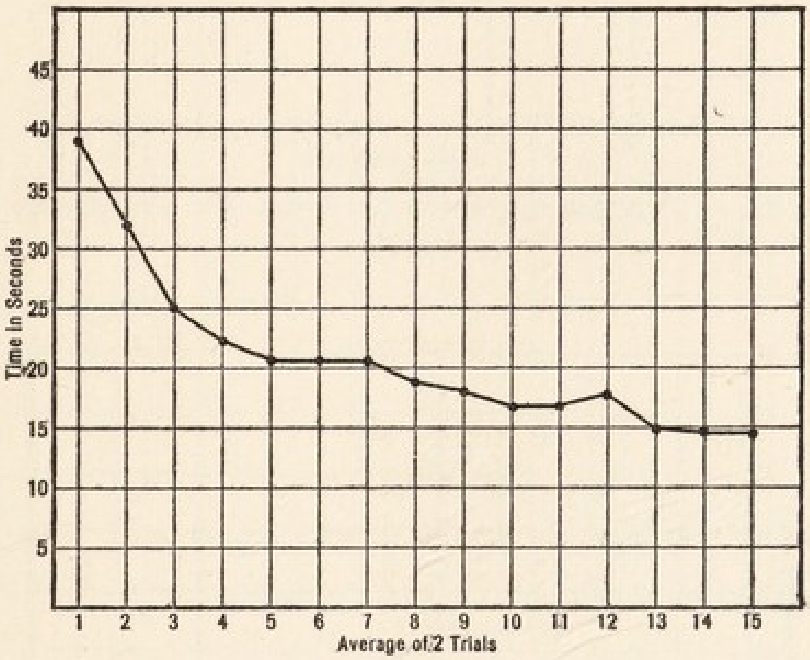


FIGURE 41. — Practice Curve as in figure 40; average of two trials.

very definitely that it is much harder to make improvement towards the end than at the beginning of the practice. Would improvement continue after the 30th trial? Very probably it would, if the subject of the experiment had continued, assuming that he retained interest in the work. We do not really know, but it is most likely. We can safely guess, however, that improvement would be more and more difficult to attain and the curve would tend to flatten out so as to become parallel with the base line.

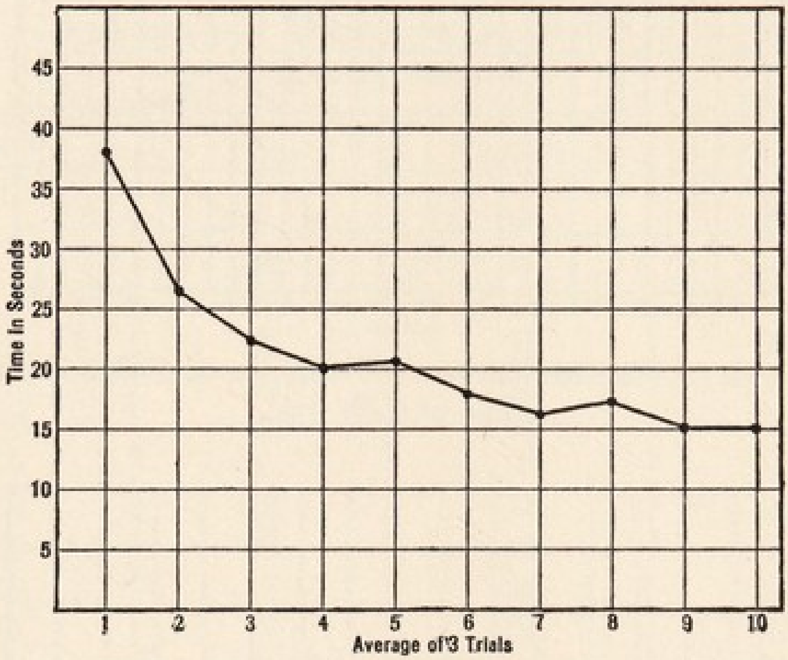


FIGURE 42. — Practice Curve as in figure 40; average of three trials.

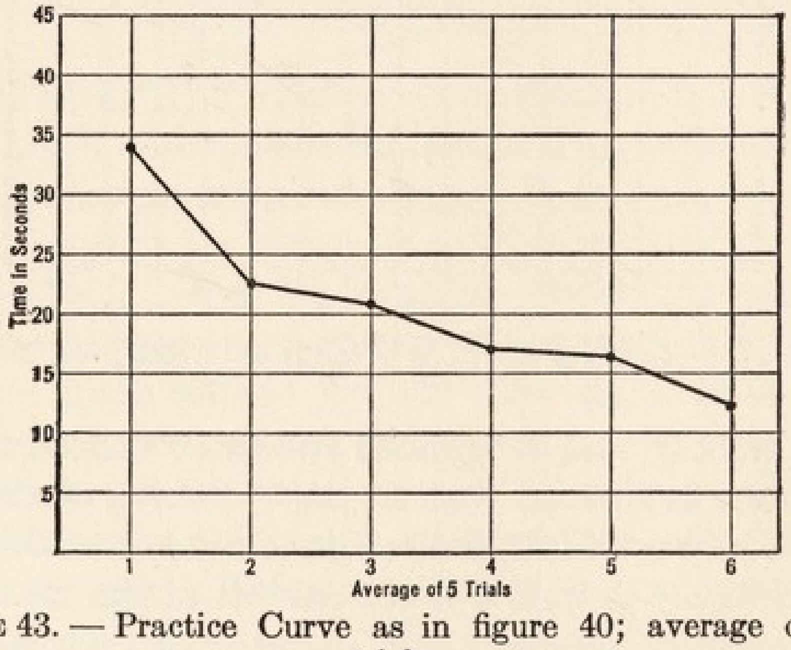


FIGURE 43. — Practice Curve as in figure 40; average of five trials.

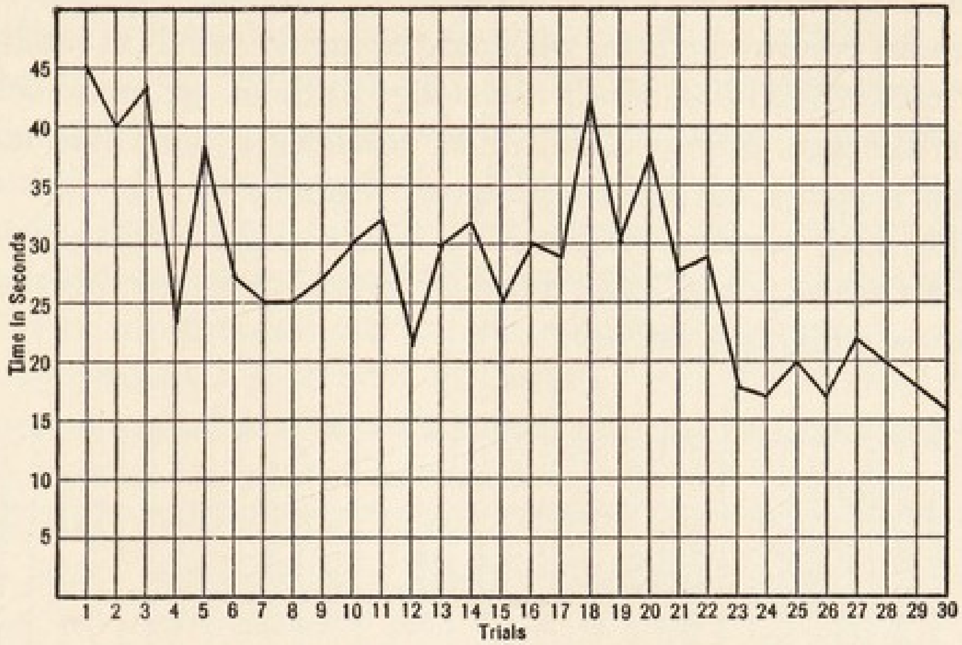


FIGURE 44. — Practice Curve. Writing the alphabet backwards. Subject A.

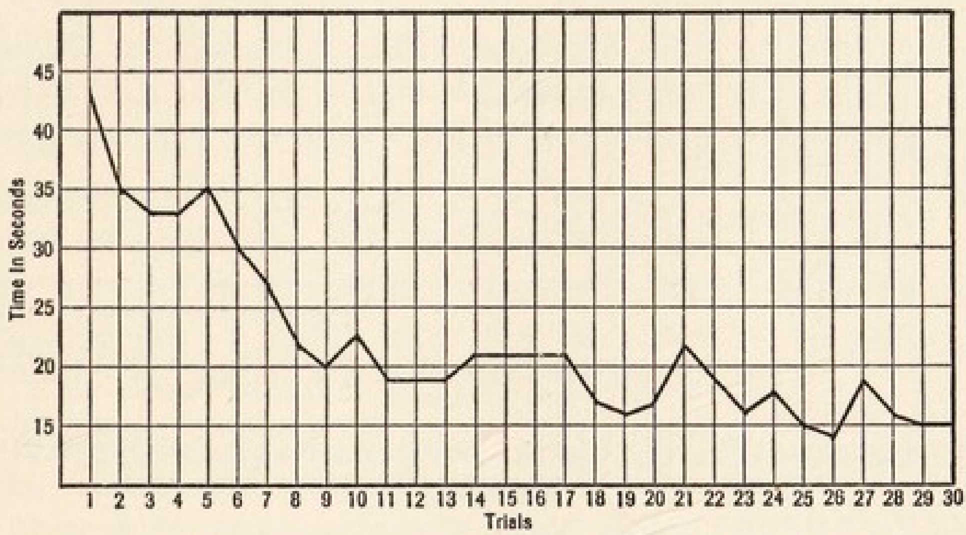


FIGURE 45. — Practice Curve. Writing the alphabet backwards. Subject B.

Figure 40 shows all the irregularities in improvement from trial to trial, and it is important to study our results in this form, so that we may be aware of the way in which learning takes place,—an irregular process from point to point, with many mistakes and many back-slidings. We may, however, also wish to consider the improvement in a more general fashion without paying so much attention to each individual trial. For this purpose we may take the average time of two, or of three or more trials and plot a curve of these averages. Figures 41, 42 and 43 show such curves. We tend to iron out the irregularities as we average a greater number of trials and we call such curves smoothed curves.

FIGURE 46. — Practice Curve. Writing the alphabet backwards. Subject C.

Smoothed Curves. — If we examine the curves we see how the irregularities slowly disappear as we proceed from Figure 40 to Figure 43. Figure 43 shows us the general trend of improvement from the first set of five trials to the second set and so on. Progress from one set to another is consistently marked by diminished time. Figure 43 shows well the general trend of the whole practice experiment, indicating marked progress during the early stages with a diminishing rate of improvement thereafter. It also suggests the possibility of further improvement, had the practice continued, because the curve is by no means parallel to the base line. The detailed curve showing every trial gives us the best picture of what actually takes place, the smoothed curve allows us to see more clearly the general trend of the whole improvement.

Individual Differences in Curves. — If the student has carried out the experiment as described in this chapter, he will be able to compare his results with the curves given here. There may be quite a difference. Each individual has his own method and rate of improvement.

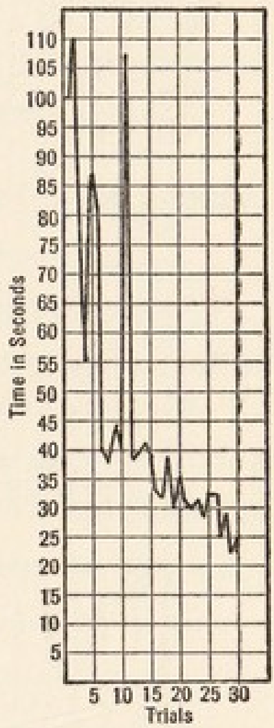


FIGURE 47. — Practice Curve. Writing the alphabet backwards. Subject D.

I have given this exercise in learning to many of my classes in educational psychology. Figures 44 to 49 show six curves for the same exercise for six different individuals. Figure 44 shows very irregular improvement with a very decided set-back before the final improvement takes place. Figure 45 shows continuous and steady improvement all the way along, a great contrast to the erratic improvement of the preceding individual. Figure 46 shows a learner who starts out very slowly and then makes very great and very rapid gains. This learner takes 250 seconds for the first trial as compared with the 45 and 43 seconds for the two preceding learners. Figure 46 also shows a very rapid increase in ability from the beginning of the learning down to trial 18 and from then on very little further improvement. Figure 47 shows a learner who starts out very poorly and who is also rather erratic. Trial number

11 is almost as bad as the first two trials. Figure 48 shows a learner who makes little improvement after the nineteenth trial. Figure 49 shows very steady improvement all the way along.

FIGURE 48. — Practice Curve. Writing the alphabet backwards. Subject E.

These curves are good examples of individual differences in learning. They are curves of university students, all of high intelligence and all mature people. Even among a homogeneous group of this kind, we find marked individual differences in learning.

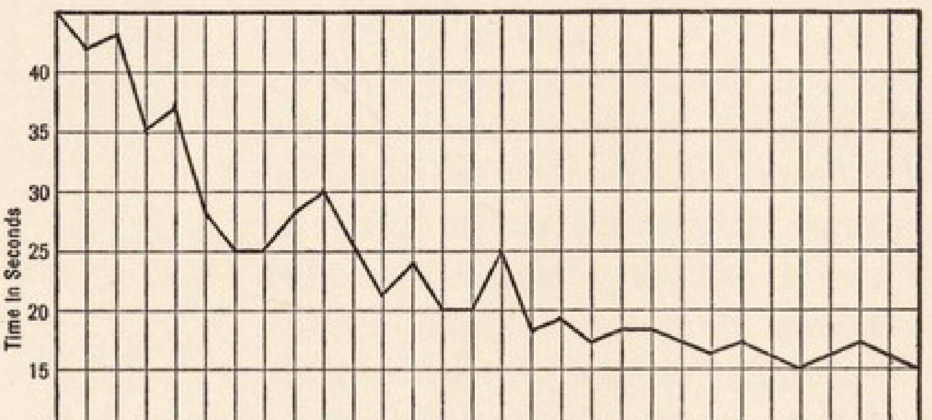


FIGURE 49. — Practice Curve. Writing the alphabet backwards. Subject F.

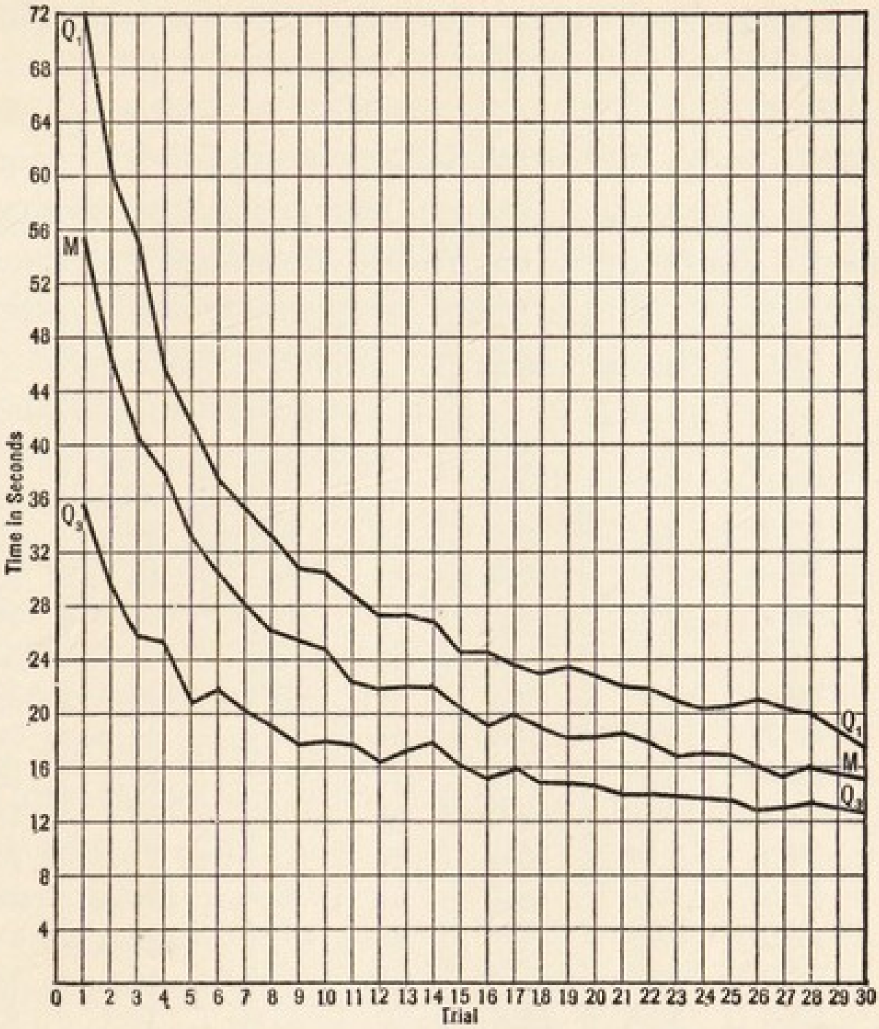


FIGURE 50. — Composite Learning Curve. Writing the alphabet backwards.

A Composite Curve. — We might ask the question as to how a university class as a group improves in the function under consideration. What is the general learning ability of the group? Figure 50 and Table 3 give the facts for a class of 88 students. The curves show the median, the upper and the lower quartile performances for each of the thirty trials. The curves are much smoother than those for any one individual. The table shows the tremendous range in individual ability from trial to trial. At the very first trial we have individuals who are able to write the alphabet backwards quicker than other individuals at the last trial. The class as a whole, however, shows steady improvement from the beginning of the practice to the end.

TABLE 3

TIME IN SECONDS FOR WRITING THE ALPHABET BACKWARDS: MEDIANS AND QUARTILES FOR EACH TRIAL FOR 88 STUDENTS

Trial No.	*Q1*	*Median*	*Q3*	*Trial No.*	*Q1*	*Median*	*Q3*
1	72.0	55.4	35.3	16	24.5	19.2	15.3
2	60.3	46.5	29.5	17	23.8	20.0	15.8
3	55.3	40.6	26.0	18	23.0	19.1	15.0
4	45.8	38.0	25.3	19	23.7	18.4	15.0
5	42.0	33.3	21.0	20	23.0	18.4	14.8
6	37.5	30.5	22.0	21	22.2	18.7	14.1
7	35.4	28.3	20.3	22	22.0	18.0	14.1
8	33.3	26.3	19.4	23	22.1	17.0	14.0
9	31.0	25.7	18.0	24	20.4	17.1	13.8
10	30.6	25.0	18.3	25	20.6	17.0	13.6
11	29.0	22.5	18.0	26	21.0	16.3	12.9
12	27.5	22.0	16.3	27	20.4	15.6	13.0
13	27.5	22.1	17.3	28	20.0	15.8	13.4
14	27.0	22.0	18.0	29	18.8	15.5	13.1
15	24.5	20.5	16.3	30	17.6	15.1	12.7

Curves of Different Functions. — So far we have been discussing the practice curve of a certain specific function, namely learning to write the alphabet backwards, and we have learned something of the general nature of practice curves, but we must not suppose that all practice takes place exactly like this. Some things are learned more slowly and laboriously than others. It would be well, therefore, to examine other practice curves in order to get some idea of the different shapes they may assume.

The curves shown in Figures 51 to 55 have all been made by students of mine as class exercises. The students themselves chose what things they would practice. Figure 51 shows the decrease in errors from day to day made by a student in using the touch system on the typewriter when copying a given list of words. The errors decrease rapidly from 16 to 4, but for the last five days the number of errors remains about 4 or 5. Evidently most of the improvement has been made during the first five days' practice. During the last five days the student makes little or no gain. This cannot be a so-called physiological limit, because it is obviously perfectly possible for the student to learn to type the list of words without error, and he would have been able to do this if he had continued at the task. He has merely reached a place in his practice where for some reason or other little improvement takes place. This makes the curve appear flat and this flat part is frequently called a plateau. Plateaus or periods of no improvement during the exercise of a function are sometimes found. They do not always occur, however. There are probably many reasons for their occurrence. We all know what it means to "go stale" at a task. We do not seem to make any progress, and we become discouraged. We are on a plateau in our learning. Sometimes our "staleness" is caused by sheer physical fatigue. Sometimes we lose interest and we don't try so hard. If we don't want to improve we don't (Law of Effect).

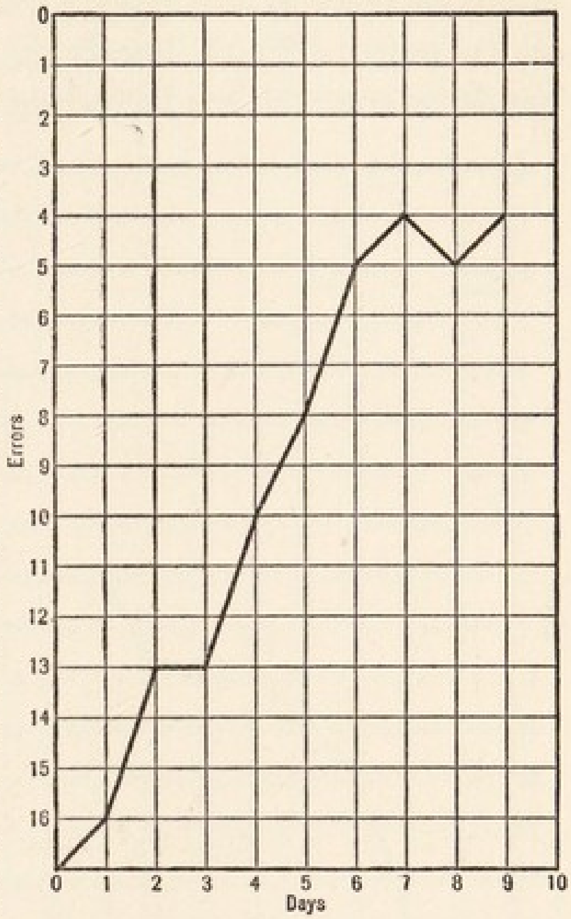


FIGURE 51. — Practice Curve for typewriting, showing the number of errors in typing a list of 35 words by the touch system: one trial a day for ten days.

Sometimes, in spite of all our trying, we still make no improvement, and then the reason probably is that we have not yet formed certain habits which are necessary before we can go on to more difficult tasks. Certain bonds or connections have to be formed before we can go on to more difficult bonds or connections. Until these bonds or connections have been well exercised the learner makes little or no improvement. When they are well formed another spurt of improvement takes place, so that if our student had continued with the task shown in Figure 51, the probability is that he would have overcome the difficulties holding him back and eventually have reached a place where no errors were made.

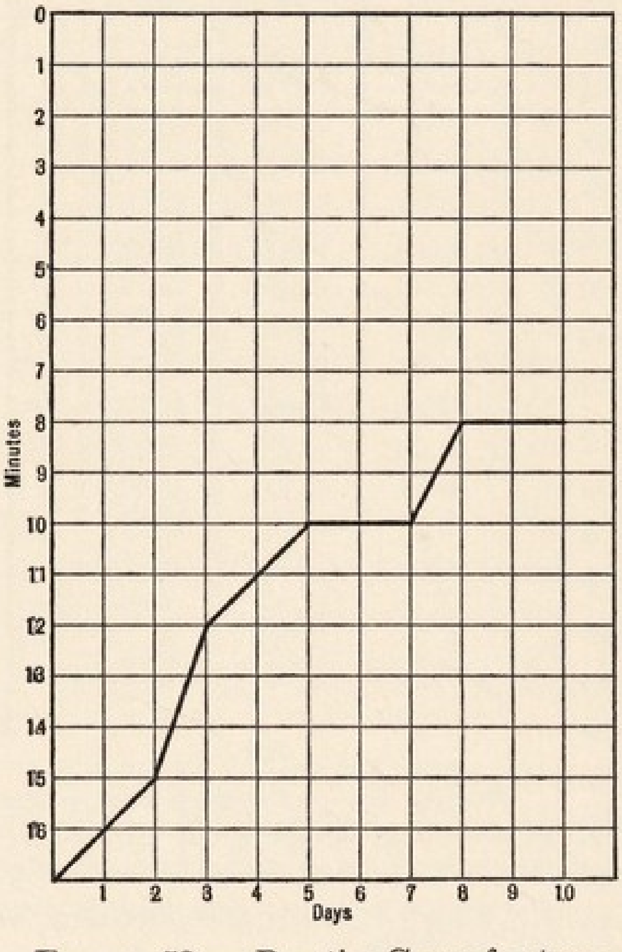


FIGURE 52. — Practice Curve for typewriting as in figure 51, showing the time in minutes for each trial.

Figure 52 shows the curve for speed in typewriting for the same student. He increases in speed fairly regularly for the ten days of practice. Notice that this student has drawn his curves differently from all the curves we have so far discussed for writing the alphabet backwards. The curves in Figures 51 and 52 go up to show gain in improvement.

Figure 53 shows the improvement in cancelling the letters B, L, V and O in a closely written magazine article.

The student worked for ten minutes each time for ten days. The curve shows the number of letters crossed out in each day's work. During the first four days improvement was relatively slow. Evidently certain habits or adjustments were being made, such as ability to keep the four letters clearly in mind. On days five and six very rapid improvement takes place. Something has been learned that facilitates the work enormously. From then on to the end of the experiment, improvement is marked, but irregular.

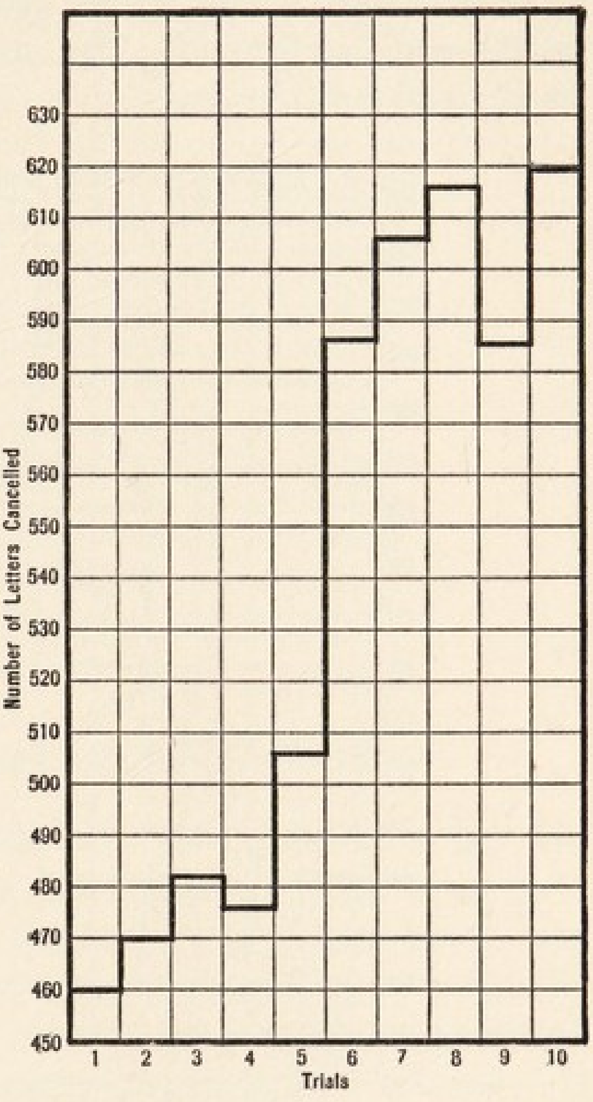


FIGURE 53. — Practice in cancelling the letters B, L, V, O in a magazine article; one trial of ten minutes per day for ten days.

Figure 54 shows improvement in substituting numbers for letters. Each letter of the alphabet is given a number and then, with the key before him the student writes the numbers for the letters of a given passage of prose. The curve shows each five minute practice period, four such periods each day for six days. The number of letters written increases during the course of practice. The curve is very irregular, because it shows each five minute period. If, however, we plot the average number of substitutions for each day's work we have the smoothed curve superimposed on the irregular curve, and this shows an almost steady and equal improvement from day to day. Here there is no sign of a plateau, no sign of going stale or losing interest or having to form certain habits before other and more complicated habits can be formed. The march of improvement is steady and continuous.

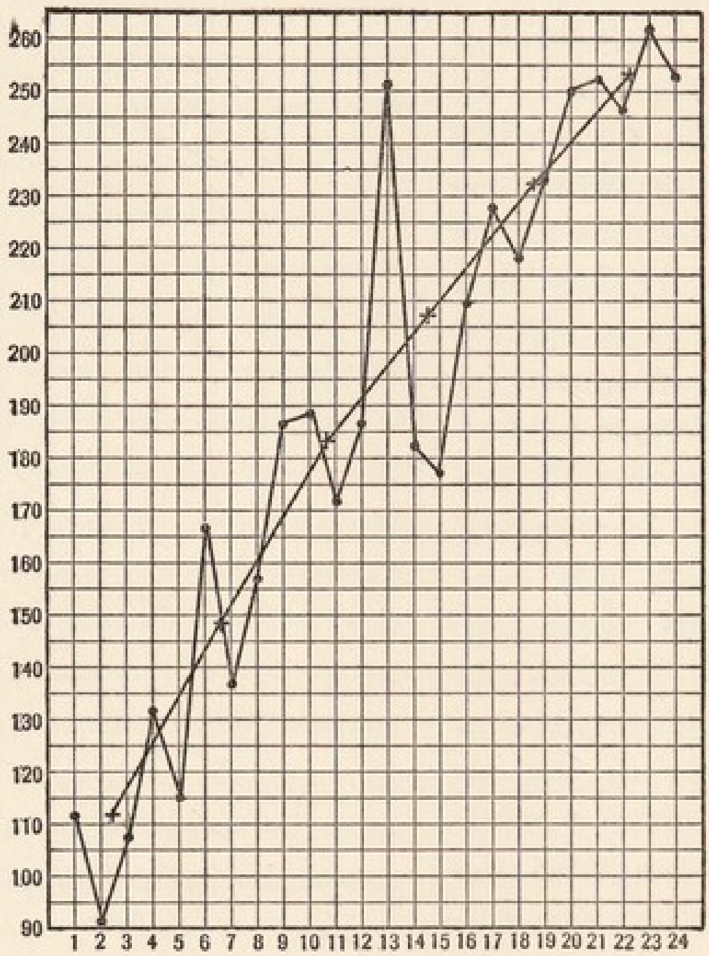


FIGURE 54. — Substitution of numbers for letters: detailed and smoothed curves.

Figure 55 shows the scores made in a certain balancing game, a toy. The object of this game was to manipulate a board with a ball on it in such a way that the ball would not fall into certain holes on the board. Scores were allotted for each hole, increasing in value for the difficulty of the hole. Each point on the curve represents the total for twenty trials per day. Improvement is shown from

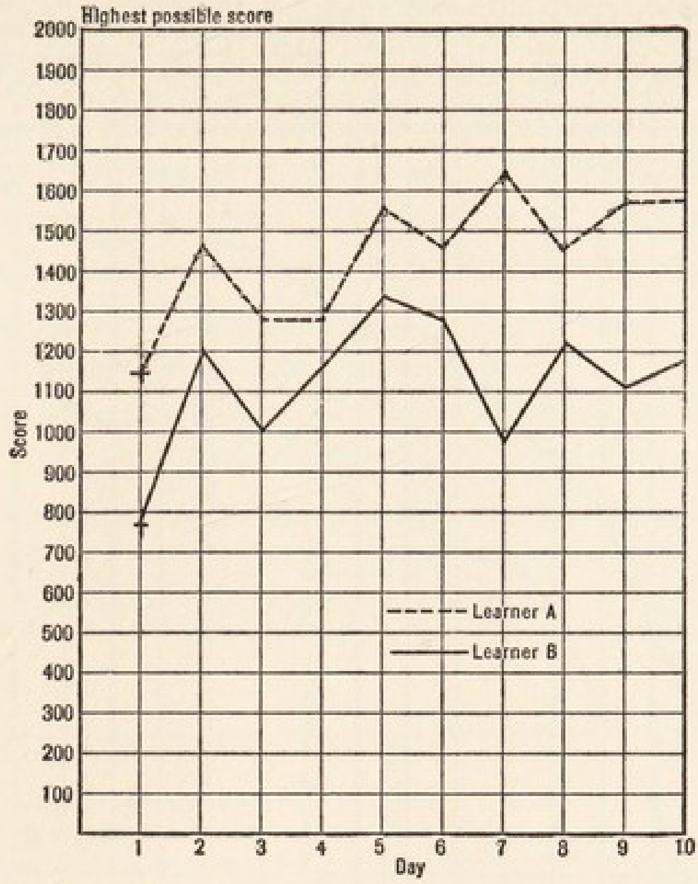


FIGURE 55. — Improvement in a balancing game.

the first to the second day, but from then on to day ten no distinct improvement takes place. Here, then, is a curve very different from most of those we have previously discussed. The learners maintained that they were equally interested during the whole period of practice, so that lack of improvement could not be due to lack of interest. It would, therefore, seem as if the function was very easy and that a physiological limit was soon reached, variations in score from day to day being thereafter due to chance. We have no proof that a physiological limit has been reached. To make certain of this a great deal more practice would be necessary.

All these curves were made by students in classes in educational psychology. The reader should plot a few curves for himself for different functions. If you are a member of a class in educational psychology, you can

DAY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	136	188	264	307	352	374	391	415	427	420	443	456	460	468	459	465	480	484	482	485
B	107	149	183	236	270	314	332	335	350	354	387	368	372	390	385	396	422	428	433	436

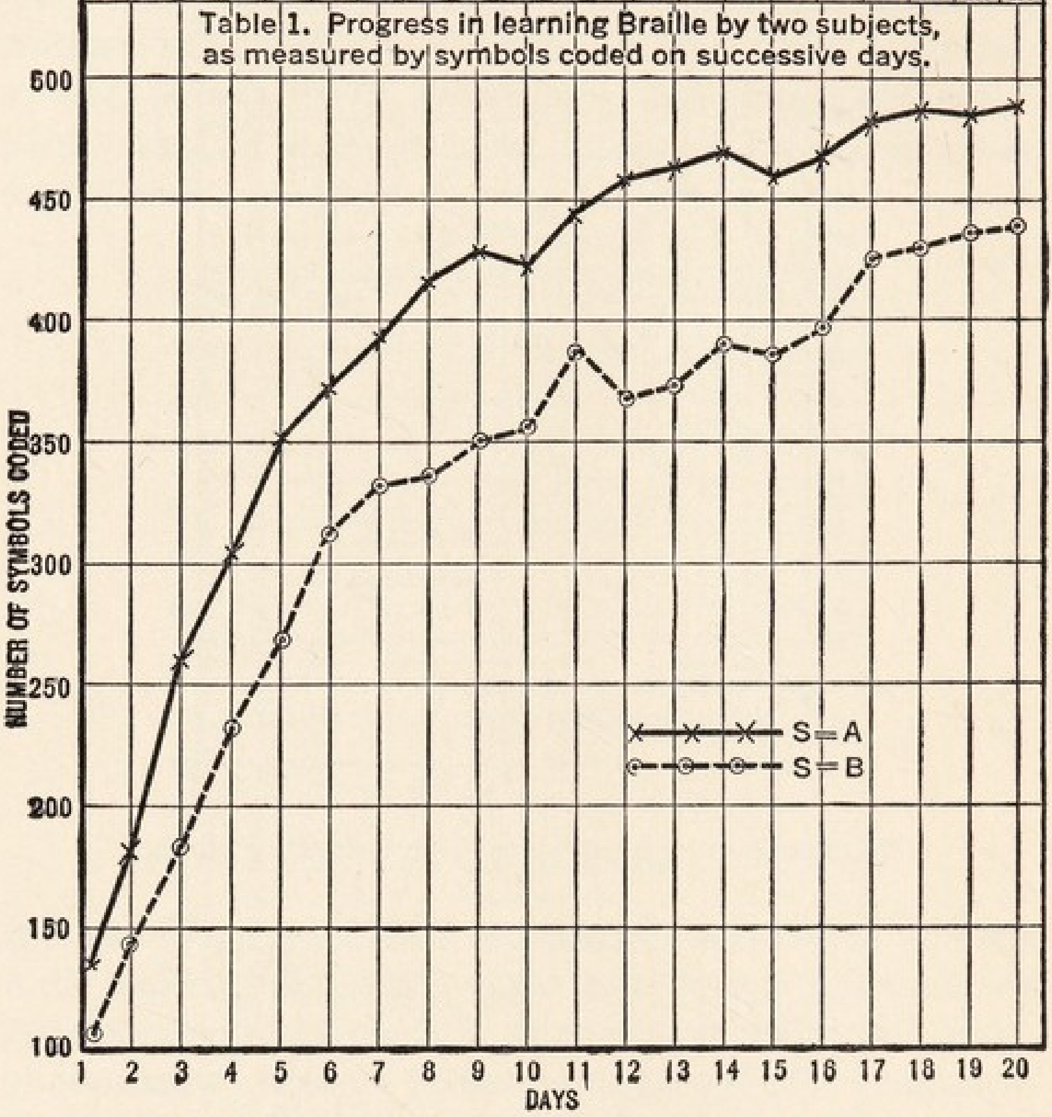
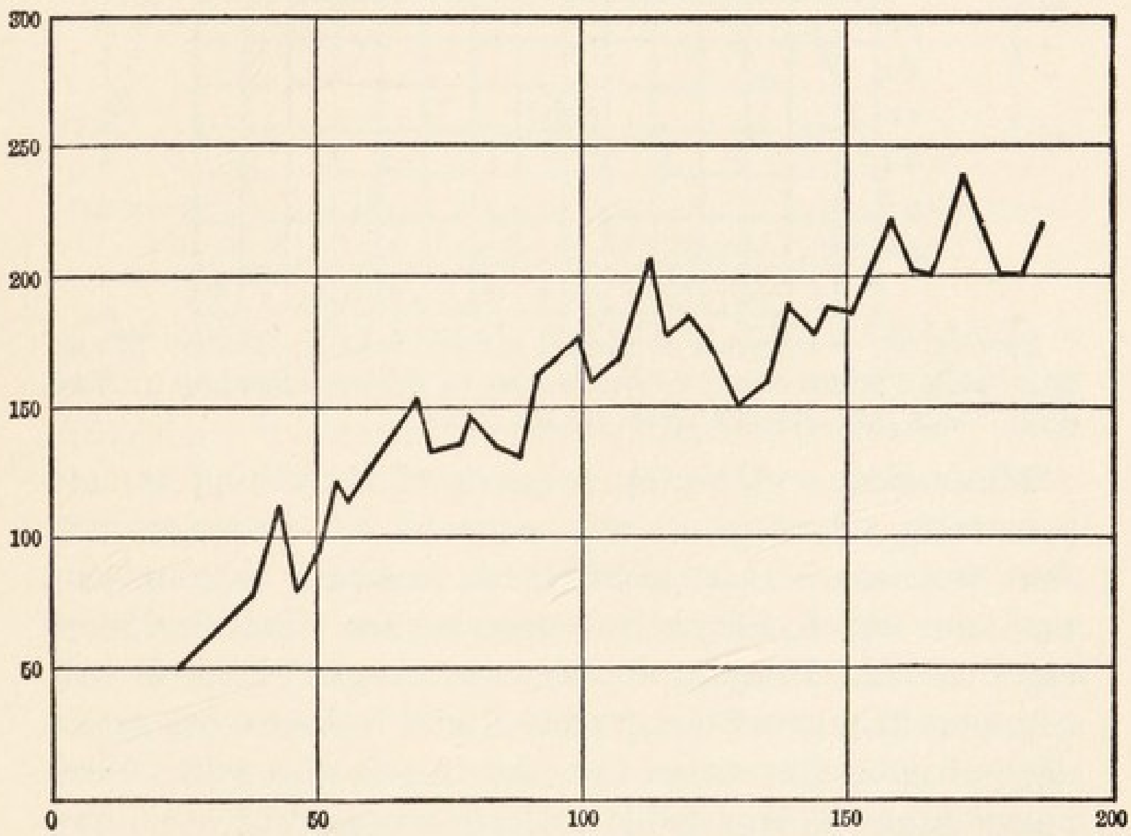


FIGURE 56. — Learning Curves of two subjects. (From Foster's *Experiments in Psychology.*)

obtain a great many curves for different functions by having each member plot a curve for a different function. In this way you can compare different functions, as well as individual differences in the same function if you have carried out the exercise described at the beginning of this chapter.

We shall now give a few curves selected from other writers' works. Figure 56 shows the learning curves for two subjects. The general tendency of the two curves is very much the same. Figure 57 shows a curve for learning typewriting. It is very irregular, but there is a gradual trend upward, showing steady improvement. Figure 58 shows the progress of a blind child, thirteen years old, in learning to read Braille. The points on the curve do not show daily tests, but tests given now and again over a period of about four months. The general trend is just about the same as in the learning of a very narrow function such as writing the alphabet backwards done all on one day. Figure 59 is particularly interesting. It shows composite curves for two groups of children, one normal group and one defective group, of equal mental age. The learning progress for the two groups is substantially the same.



Courtesy of the Journal of Applied Psychology

FIGURE 57. — Learning Typewriting. The vertical axis represents the score in a five-minute test in typewriting. The horizontal axis represents the number of hours of practice. (From Chapman, *Journal of Applied Psychology*, 1919.)

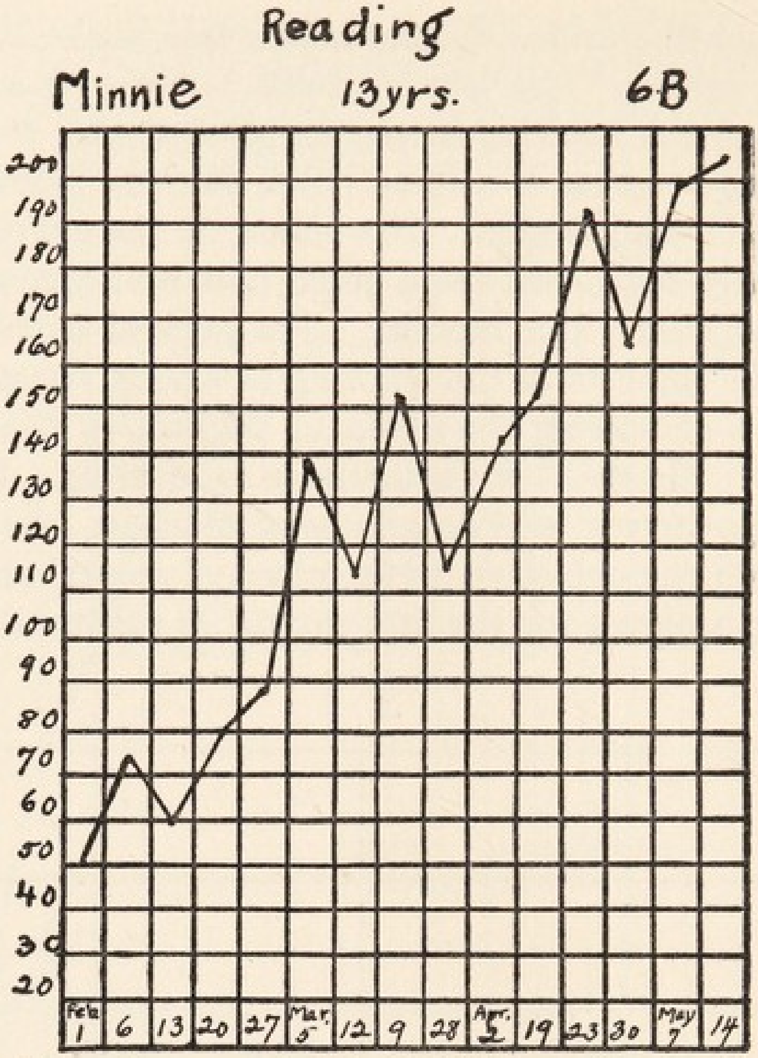


FIGURE 58. — Progress of a blind child in reading Braille. (From Maxfield's *Present Status of Instruction in Primary Reading in Residential and Day School Classes for the Blind.*)

Motivation. — Drawing a curve of something we are practicing shows us at any moment our proficiency in that function. It is much more accurate than relying upon our own feeling of how good we are. It is like keeping a record. Keeping records of one's proficiency is very common in a great many games, and indeed most games depend upon the record in order to tell who won. With many games it is possible to keep a record of one's own score in order to see whether one is improving or not, and trying to beat one's own record becomes just as exciting as beating one's opponent. The golfer is just as anxious to lower his record as to beat his opponent.

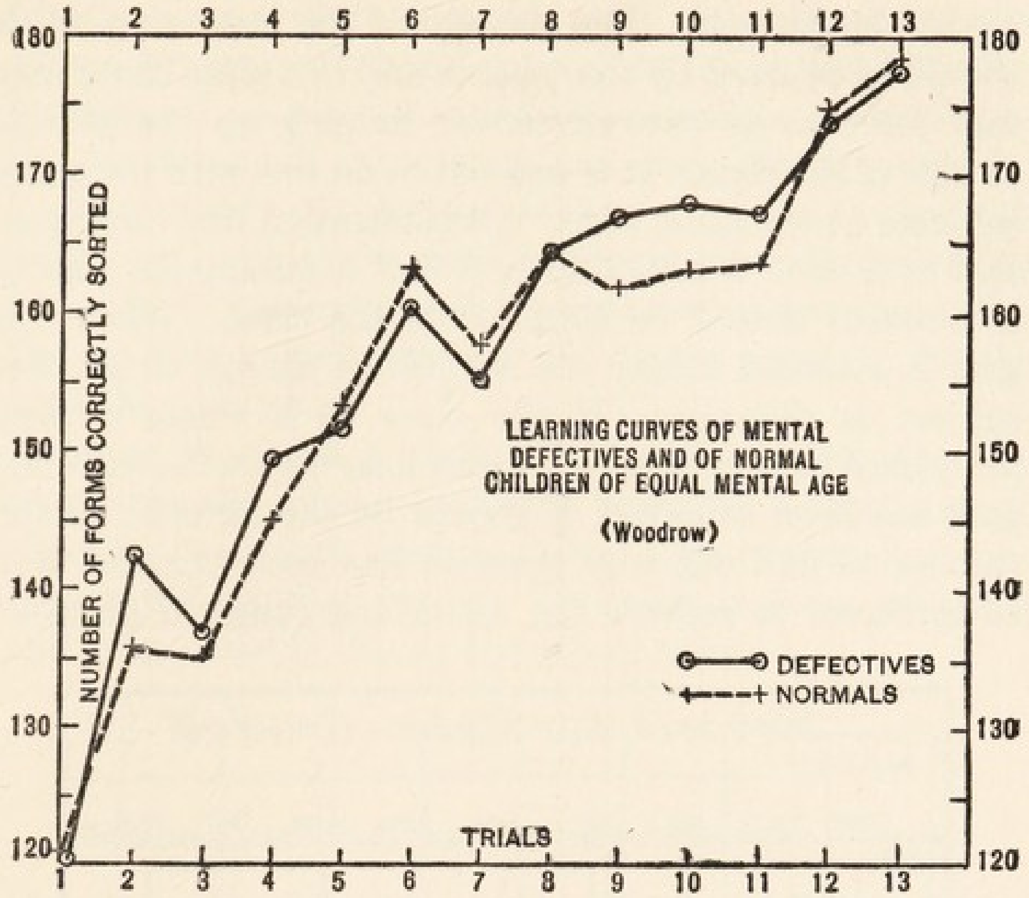


FIGURE 59. — Learning Curves of Normal and Defective Children. (From Woodrow, *Journal of Educational Psychology*.)

One reason why keeping records in intellectual tasks is not as common as in physical tasks is probably that it is not as easy to measure the former as it is the latter. Now that we have measuring devices for all the school subjects, many teachers are using them to measure the progress of their pupils. Children should be taught how to keep such records of their own progress and if this is not overdone they will find a great incentive to improvement by keeping such records. They like the actual making of the curves themselves and they experience satisfaction as the curve ascends, and satisfaction as we have already learned is the potent factor in the law of effect. It helps to strengthen the right connections. Standard tests of arithmetic, writing or reading may be given every week or every month and the progress of each individual recorded by himself. The average of the class as a whole should be charted by the teacher and this class chart may well serve as an instrument for keeping up the general morale of the class. It is well not to do this with too many subjects at the same time. Concentration upon one subject at a time is probably best and a reasonable goal of attainment should be kept before the class. When this goal is attained in any one subject, a change to another subject is desirable. If the class as a whole or any individual child reaches a plateau long before the expected goal has been achieved, it should be the function of the teacher to find out why progress has been arrested, and to endeavor to remove the disturbing causes if possible.

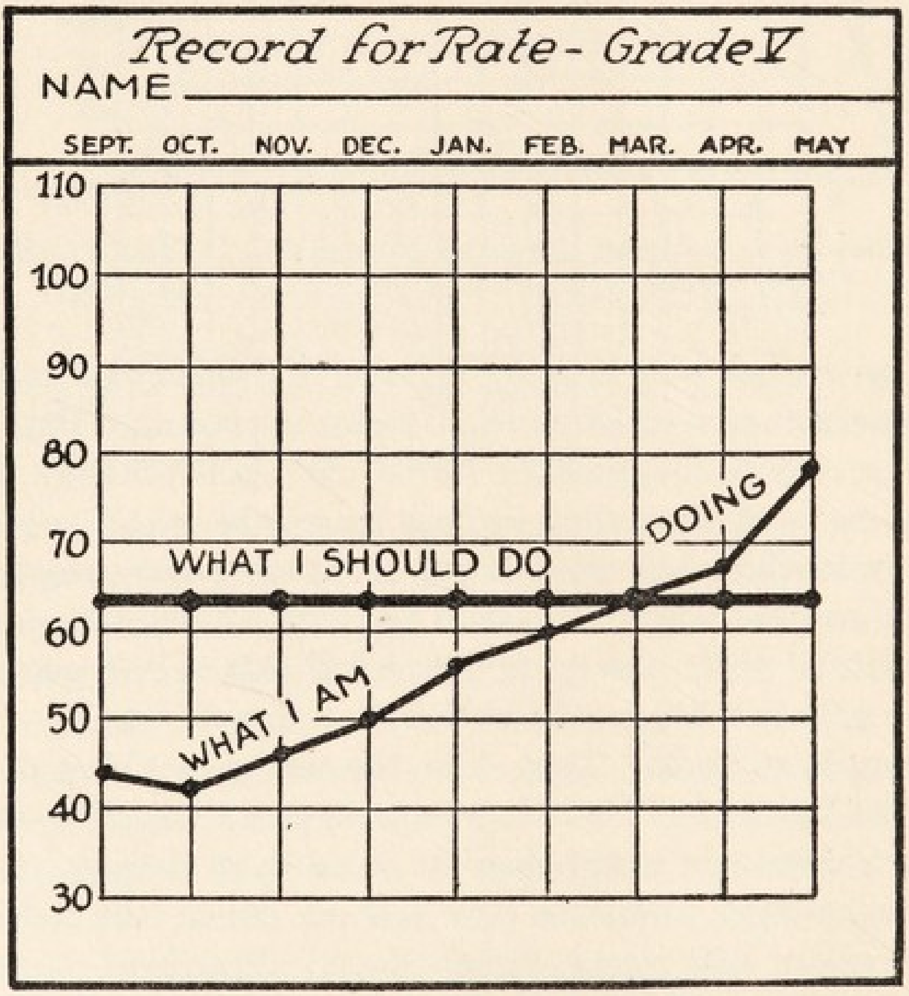


FIGURE 60. — Motivation in Penmanship. (From Paulu's *Diagnostic Testing and Remedial Teaching.*)

Plateaus are seldom necessary and are most frequently caused by lack of interest, or a decrease in desire for improvement. A plateau in general should be a warning to the teacher that something has gone wrong with the learning ability of the class. Figure 60 shows the individual record kept by the pupil for the purpose of motivating his work. This chart has been used for handwriting. The line marked "what I should do" is the standard for the grade in question.

Figure 61 shows a class record. The continuous line represents the average monthly score for a fourth grade class in the Starch reading tests. The broken line represents the standard scores. This class is above standard and gains steadily during the school year.

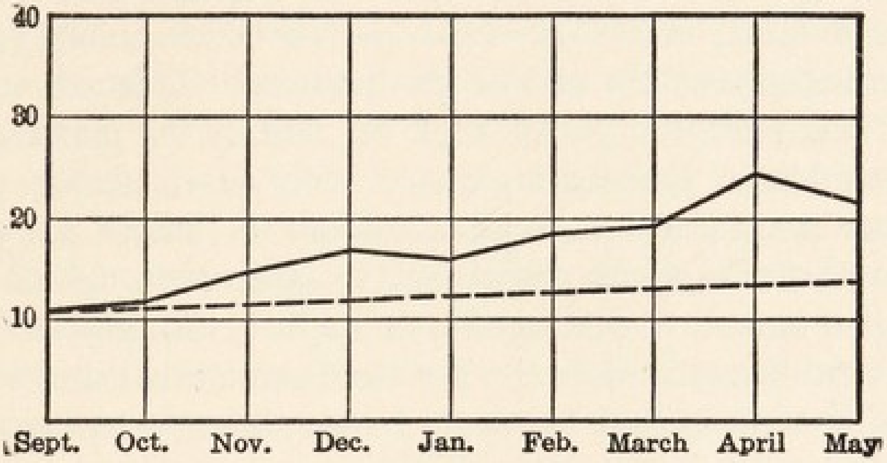


FIGURE 61. — Monthly progress in reading. The dotted line represents standard scores. (From Starch's *Educational Psychology*.)

Characteristics of Learning Curves. — Are there any characteristics which are common to all learning curves? In general the answer to this question is "no." Apart from the fact that all show an increase in the ability to perform the function in question and that all would, if indefinitely continued, reach a limit of improvement sometime or other, there are no constant characteristics to be met with in all curves. Learning curves may be of all kinds, some showing rapid initial improvement and others slow initial improvement; some show steady persistent gains, others irregular fluctuation from point to point; some have very definite plateaus, others show no indication of a plateau. Such differences depend both upon the function practiced and upon the individual practicing. It is therefore useless and undesirable to talk of a typical learning curve.

Initial Spurt. — Some curves show what has been called an initial spurt which means a period of relatively great improvement appearing at the beginning of the practice, but many curves do not show this. If it occurs, it is probably due either to the ease of forming certain initial habits, or else to the greater interest and enthusiasm of the learner at the beginning of the learning.

End Spurt. — Some curves show what has been called an end spurt, which means a period of relatively great improvement at the end of the learning. This can occur only if the learner knows that the end of the learning is approaching. Sometimes under such circumstances the learner makes a last vigorous effort to better his performance, and this is registered by a decided rise of the practice curve. Initial spurt, end spurt, and plateau are all useful terms to describe certain characteristics of some learning curves, but we must be careful not to imagine that they always occur in all curves.

Physiological Limit. — A physiological limit would occur in all curves if they were indefinitely continued. By a physiological limit is meant that stage of proficiency which cannot be surpassed owing to the physiological limitations of the organism itself. True physiological limits are probably rarely attained by anyone in any function. We do not go on practicing things until our physiological limit is reached. To do so would mean very persistent and continuous work. We have so many other things to learn in life that we have not time to devote to reaching our physiological limit in any particular function. And more important than this, we lack any incentive to further improvement in most things, because in the thousands of daily habits we have reached a level of proficiency with which we are satisfied, and so long as we are satisfied with certain responses we continue to repeat them when the appropriate stimuli recur. I am satisfied with my present proficiency in writing. It is good enough. It serves its purpose. There is, no doubt, great room for improvement both in speed and quality, but so long as I feel content with it as it is, no improvement will take place, regardless of the number of times I exercise the function. I have not attained any physiological limit in writing, because I could easily set myself the task of improving and if I thought this worth while, my writing would immediately improve in speed and quality. And so it is with rate of reading, ability to add, subtract, multiply and divide, with my efficiency in dressing, in using a lawn mower, in tending the furnace, in running a motor car, in skimming over a newspaper, in typing, in reading a novel, in peeling an apple, and all the other thousands of habits which I have formed and brought up to a certain degree of proficiency satisfactory to me at the present time. I do these things well enough for my purposes, but certainly not as well as I could, were I to concentrate on any one of them. None of them is near my physiological limit.

Physiological limits are probably reached or nearly reached in some functions by some few people who set themselves very deliberately the task of excelling in some particular thing. People who compete for championships of various sorts are probably near their physiological limits, as well as other people striving hard to excel in tasks for which no records are kept. The championship records for running and other physical exercises, the records for typing and chess and so on represent the approximate physiological limits for these functions for the human race. We say approximate limits, because we see the present records every now and then surpassed. With the ordinary school tasks, however, the teacher may feel pretty sure that children are nowhere near their physiological limits. Furthermore, the teacher has no real concern with physiological limits, because his task is to bring his pupils to a reasonable degree of proficiency rather than to strive for a limit.

Summary

1. A learning curve shows the rate and amount of improvement in the function measured.
2. Even in such a simple function as writing the alphabet backwards every learner shows marked improvement in a few trials.
3. Curves may be drawn to show all the details of improvement or smoothed curves may be made to show the general trend of improvement.
4. There are very decided individual differences in the learning curves of university students practicing the same task.
5. A composite curve showing the average performance of each trial for a group of individuals may be drawn. This shows the general trend of improvement for the group as a whole.
6. Samples of curves of different functions show that learning curves differ widely from function to function as well as from individual to individual.
7. Learning curves may also be constructed to show the general improvement in any function over a long period of time by taking sample measures of this function once a week or once a month.
8. The learning curves for defective and normal children of the same mental age seem to be very similar.
9. Keeping a record of one's proficiency in a subject is an excellent method for sustaining one's interest in the work.
10. Individual and class curves of school work may be used to great advantage as a means for motivation in school.
11. There are no general characteristics common to all learning curves.

12. Most learning curves are irregular, showing that improvement takes place unevenly.
13. A plateau is a period of little or no improvement, followed by improvement later on.
14. Plateaus may be caused by lack of interest or incentive, "going stale," fatigue or the necessity for the formation of simpler habits before more complicated habits can be formed.
15. Plateaus are not commonly found in learning curves.
16. A plateau should be a warning to the teacher that something has gone wrong with the learning ability of the class.
17. Initial spurt and "warming up" are names applied to the preliminary improvement found in some curves.
18. End spurt is a name applied to a sudden increase in improvement found at the end of some learning curves.
19. End spurts do not occur unless the end is known by the learner himself to be near.
20. A physiological limit is the highest point on the curve that the learner can reach.
21. Few of us ever achieve a physiological limit in any function.

Review

True-False Statements

As a review mark the following statements true or false:

1. Learning curves always show regular and constant acceleration from one trial to the next.
2. Most subjects in learning to write the alphabet backwards arrive at a plateau at about the seventh or eighth trial.
3. A rapid initial rise in a learning curve is universal.
4. A rapid initial rise is sometimes called an initial spurt.
5. There is no one typical curve of learning.
6. Most of us remain on plateaus in regard to the greater number of our daily habits.
7. If a child reaches a plateau on his learning curve for addition it is an indication to his teacher that he is becoming lazy.
8. To have children keep records of their improvement in different subjects is a very poor incentive for them to strive to do better.
9. University students practicing the same task are likely to show very similar learning curves.
10. The most common characteristic of learning curves is their irregularity.
11. Almost every learning curve shows a plateau.
12. When a curve suddenly accelerates for the last two or three trials, we call this acceleration an end spurt.
13. The learning curve differs according to the person who does the learning.
14. Plateaus are more likely to occur in learning complex tasks rather than in learning simple tasks.
15. The plateau in a learning curve indicates a period of improvement.
16. A great deal of our habitual activity is not as efficient as it could be, because we are satisfied with it as it is.
17. The learning curve gives a graphic picture of the amount, rate, and limit of improvement.
18. We soon reach our limits of efficiency in most tasks because our organism is so unmodifiable.

19. Underline the best response:
A long period of no improvement followed by further improvement is called: (1) a plateau; (2) adaptation; (3) initial spurt; (4) warming up.
20. Check the best response:
We do not reach our physiological limit on most tasks, because:
(a) of our lack of intelligence.
(b) our learning tends to decline.
(c) we are satisfied with our present level.
(d) of lack of sufficient practice.

Advised Reading

Thorndike, E. L.: *Educational Psychology: Briefer Course*, Teachers College, Columbia University, New York, 1914, Chapters 14 and 16.

CHAPTER X

EFFICIENT LEARNING

Effective Ways of Learning. — So far we have discussed the general laws of learning and the general characteristics of improvement in simple functions. We may now ask the question as to how such knowledge may help us to learn more effectively. In this chapter we shall discuss some of the consequences of the general principles of learning and in addition add what knowledge we possess as to effective learning gathered from experimental work in this field. There are thousands of problems with respect to the best means of learning and many specific points have not yet been definitely settled by experiment. Nevertheless, there is a growing mass of knowledge resulting from reliable experiment in the learning process, and we will discuss here the most useful facts for the teacher and learner.

The Learning of Arithmetic. — One of the most effective methods of approach is to ask what help educational psychology can give to the learning of a particular subject. Thorndike does this very effectively in his "Psychology of Arithmetic." The problem with reference to the learning of any school subject is, he says, "the development of a hierarchy of intellectual habits." What we have to do is to build up a great many habits. Many desirable habits are very complex. We cannot form them right away. We need other simpler habits first in order that the more complex ones shall be able to function. For instance, we want our children to learn to calculate interest on money saved. We cannot teach them this right away, because it involves knowing how to add, multiply, and so on. Hence the necessity for forming simpler habits first. That is, as Thorndike says, we must build up a series or hierarchy of habits. In order to do this, three important considerations emerge, according to Thorndike:

1. The choice of the habits to be formed.
2. The best order in which to form them.
3. The best means for forming each in that order.

The first point means that we must go over all the details of a subject and seriously consider just what we want the child to learn. And, further, we must decide what habits he must form in such learning. It is easy to say that we want the child to be able to add 2 and 3 together. We must go further and decide whether on presentation of the stimulus 2+3 or $\begin{array}{c}2\\ \underline{3}\end{array}$ we want the child to respond by saying to himself "2 and 1 are 3 and 1 are 4 and 1 are 5," or "two, three, four, five" or "two and three are five" or simply "five." Obviously the last response is the most economical, namely to think "five" on presentation of the stimulus 2+3. It is in this detailed sense that we must determine what particular habits should be formed.

When we have made such decisions, we must then find the best order in which to form these habits. Thorndike's rules for arithmetical bonds or habits are as follows:

1. Other things being equal, one new set of bonds should not be started until the previous set is fairly established, and two different sets should not be started at once.

As an example, first learn multiplication without carrying, then with carrying and no zero difficulties, then introduce the zero. Or in teaching punctuation, teach the use of periods and commas first, then the use of question and quotation marks and finally the use of colons and semi-colons.

2. Other things being equal, bonds should be formed in such order that none will have to be broken later.

It is probably better to teach long division with remainders from the very start, rather than to teach it without remainders first. The child will then not form the bond of "coming out even," which must later be broken. In reading, diacritical marks are of doubtful value because bonds are formed, which later on have to be broken. Some teachers of public speaking believe that students should be taught right away to speak without notes. If they learn to speak first with notes, they find it difficult to break this habit later on.

3. Other things being equal, arrange to have variety.

Do not work on addition until all the necessary bonds have been formed. After some have been fairly well established, move on to subtraction, and then to multiplication.

4. Other things being equal, use objective aids to verify an arithmetical process or inference after it is made, as well as to provoke it.

After a pupil has responded to $2+3$ by 5, he may check his work by actually counting three objects and two objects placed together in a pile. This will serve as a concrete check upon what he has already done abstractly.

5. Other things being equal, reserve all explanations of why a process must be right until the pupils can use the process accurately, and have verified the fact that it is right.

A child can profitably learn how to add, subtract, multiply, and divide with United States money long before he knows anything about decimals, or the reason for placing decimal points where we do in such operations. The important point is to be able to get the right answer and be able to verify it. The reasons for the particular method of procedure will be appreciated better after the procedure has been learned.

6. Arrange the order of bonds with due regard for the aims of the other studies of the curriculum and the practical needs of the pupil outside of school.

This is obviously a general rule to keep in mind with all studies so that they may be linked together and function in the life of the pupil.

These are some of the general principles which Thorndike uses in his investigation of the most effective way of learning arithmetic. Everyone of these and other principles is illustrated by numerous examples. The student should read the whole book as the best example of the application of the psychology of learning to a particular subject.

Learning to Read. — The investigations of Gates (1 and 2) in reading may be taken as another good example of the application of psychology to a particular subject. Gates started with a study of the reading disabilities of pupils of all ages. In order to understand these reading difficulties, he constructed many kinds of reading tests. These tests help to diagnose the specific difficulties of each pupil. Hence we have a battery of valuable tests measuring the many specific abilities which go to make up the complex act of reading.

This work naturally led this psychologist to raise the question as to whether reading might not be so taught from the beginning as to avoid the difficulties found in so many pupils later in life. Hence he made an analysis of the ways in which young children learn to read without specific teaching. He found that in the recognition of words, different children seized upon different characteristics in order to enable them to remember the word, thus "monkey" was remembered by the monkey's tail on the *y;* or by the "hole," that is the letter *o* or by the "funny chair," that is the letter *k* and so on. Furthermore words were differentiated by means of their length as "cow" and "dandelion." In other words, children left alone by themselves use all sorts of devices, that is they form all kinds of habits. Now some of these habits are useful and help greatly, but others are useless and may retard the learning. Hence we need to analyze out the most useful habits and teach those, and at the same time not allow the useless and wasteful habits to be formed.

In these investigations Gates found that many teaching devices were training the children in bad habits, that later on would have to be discarded. Examples of such wasteful devices are the use of flash cards and numerous phonetic devices used in the so-called phonetic method of teaching. All these devices he calls extrinsic, i.e. devices that must be learned in addition to the reading proper. Hence he discards them and in his proposed method of teaching reading he uses the intrinsic device only. The intrinsic device is a help and aid to learning some specific aspect of reading, but at the same time is an integral part of the whole reading process. It is not something that has to be learned and cast away later on. As an example of an intrinsic device to make children "observe closely certain characteristic parts of words and gradually to become aware of certain elements common to many different word-forms," the pupil is given a drawing with the following instructions:

1. Color the cat blue.
2. Color the hat black.
3. Color the coat brown.

and so on with many similar exercises.

This, of course, merely gives a bare outline of Gates' work. It is, however, a good example of a teaching method which has been arrived at by a psychological analysis of the learning process.

Learning is Reacting. — Since all learning consists of reacting to various stimuli, we should as far as possible see to it that during our learning we are reacting in the same way as we later wish to react, on the presentation of a specific stimulus. This is so obvious in motor learning as in learning acts of skill that we do not need to elaborate. We learn to drive a motor car by driving one, not by studying a diagram. Of course the diagram may help in certain ways, but we should certainly never think of studying all the moves on paper and then expect to be able to get into a car and start out driving perfectly, however perfect our "paper" driving may have been. Obvious as this is in acts of skill, we frequently forget it in other kinds of learning. We learn grammar in order to write or speak correctly. But we do not find that we write or speak correctly, whereas we do know more grammar. If we want to write and speak correctly, we must practice writing and speaking. Study your French lesson "out loud" if you want to learn to speak French. Keep writing poems if you want to learn to write poetry.

If you are going to be examined on a subject and you wish to do well in the examination, consider how the questions will be asked and practice answering in the same manner that you will be called upon to react. If it is an oral examination frame possible questions and talk them out aloud *extempore* so that you can judge for yourself how they will sound to the examiner. If it is an old-type written examination, write down answers to possible questions in essay form. If it is a new-type examination construct true-false, alternative answer and completion items and after several days answer them yourself. You will in all probability in this last case, get all your answers right, because you will find that the construction of objective items is an excellent form of review.

We learn exactly what we practice. A great deal of the controversy over the value of different methods of teaching arises from the fact that we frequently forget this rule — that we learn exactly what we practice. Is one method better than another? Better for what purpose, must be our first consideration before we can answer the question satisfactorily. Is the lecture or laboratory method better? If our aim is the acquisition of a certain number of facts, the lecture method will probably achieve this result more economically. If our aim is to acquire a certain amount of skill in solving the problems of a subject, the laboratory method will be most effective. To become efficient mental testers the laboratory method of teaching is necessary, to acquire a knowledge of the development of mental testing and its present significance and value in education, the lecture method would be quite adequate.

Models are a help in learning just because they help to check up our responses and let us know how far off we are from the desired response. But models should not be so perfect, as to make it hopeless for the child ever to attain them. This was the fault of the old copy-book models in writing. They were too far beyond any possibility of attainment by the child.

Reaction of some sort is necessary in all learning. "React—react," must be the watchword of all learners. Passivity is waste of time. We do not learn by absorption. Sitting quietly gazing at the book as so many children do is waste of time. Write out the spelling words or the foreign words to be learned. If you do not know what to write for your theme, get pencil and paper and write anything. You may have to throw away a lot later, but the very reactions of writing will probably set you on the right road in time.

Avoid Irrelevant Reactions. — What has just been said about reacting in the way we will need our information or skill later implies in a negative way the principle that we should as far as possible avoid making reactions which we shall not need later. Sometimes this cannot be avoided, but so far as possible, it is a safe principle for our guidance. Sometimes the easiest road seems to be a round-about way, but it very rarely, if ever, is. The teaching of such a complicated task as writing to young children has led teachers to seek all kinds of aids in the task. Many methods of preliminary practice before writing are advocated, such as tracing letters in grooves, tracing on transparent paper, running the finger over sand paper letters, practicing parts of letters and so forth. It has been demonstrated (Hertzberg, 3) that children who are taught to write without any such preliminary aids make faster progress than children who are taught by means of them. Those who begin by tracing the letters learn to trace and they have to forget certain habits learned in tracing, when they come to writing without any aid. All aids in learning are not to be condemned outright, but whenever the required reaction can itself be directly learned from the very beginning, the chances are that it is more economical in the long run to begin with it.

In the teaching of arithmetic we find many examples of irrelevant reactions. Adding by counting on the fingers is a gross example, which few teachers now-a-days would teach, yet at the same time they may feel that the child is aided by writing down the number to be carried. But this writing down the number to be carried is an irrelevant reaction which must later be got rid of and it is almost as doubtful a teaching device as counting on the fingers. These indirect methods of teaching, which are supposed to help the child, are continually cropping up. Here is a method of addition that I came across recently in a very good school.

ADD

```
     5 6 9 4 3
     8 2 1 6 9
     3 2 1 5 0
     6 6 5 8 6
     2 5 9 8 3
     1 5 8 1 6
     2 3 5 6 1
     1 4 8 1 9
     ---------
           3 7
         4 2
       5 0
     3 8
   3 1
   -----------
   3 1 8 0 2 7
```

It is of doubtful value, especially if children have once been taught the ordinary method. Some children arrived at the answer 3 5 3 4 5 7, obtained by adding all the totals, showing how other habits were functioning. But even if there are no old habits to interfere, we must remember that later we shall have to break down the habit of writing down the full total for each column and teach the child to "carry" mentally. Present-day psychology would recommend that he be taught to carry mentally from the very beginning of his learning to add, so that no useless habits will have to be discarded later.

Historically a good example of the gradual abandonment of preliminary aids in favor of direct learning is shown in reading. It used to be the thing to start learning the A B C's. What the child learned was to say A B C in sequence and to recognize the letters. Then came da, ma, pa, ta, fa, and countless syllables. If the child had learned his previous lesson properly, he would respond on seeing "da" by saying "dee-a." So the poor child had now to forget that A called for the response A, and D for the response Dee, and learn that d and a together were called *da* or *day* or something like that and so the drill in syllables would go on da, ma, pa, and me, be, te, le, and so on. Then came real words and so when "date" was presented the child would likely respond by "da-tee" as he had been taught in syllables. So these bonds had to be all broken again and much that had been painfully learned had to be forgotten. To-day we find it far more economical to begin with real words or phrases or sentences and form direct bonds with these, bonds which will not have to be broken later.

Mnemonic Devices. — The whole question of the value of mnemonic devices in learning can be best solved from this general standpoint. If the device helps in forming a real link between the material to be learned it is probably helpful, but if the device calls for the formation of a lot of useless bonds, which are in themselves hard to form, such a device is probably very uneconomical.

My schooling dates from the time when it was considered the proper thing to learn a great many facts in such subjects as geography and history. My first contact with a mnemonic device is still vividly remembered. The teacher wrote on the blackboard without any previous warning or explanation the following words one under the other, each with a capital:

On
Reaching
Eton
We
Lunched
Granley
With
Grandpapa.

Now this procedure with this peculiar sentence with the peculiar name "Granley" was all very strange and therefore very interesting. It caught and held the attention of the class. Then he told us that this sentence would help us to remember the towns on the River Thames in their order from the source to the mouth, the first letter of each word being the same as the first letter of the town. And so the mnemonic device was completed in this fashion:

On	Oxford
Reaching	Reading
Eton	Eton
We	Windsor
Lunched	London
Granley	Greenwich
With	Woolwich
Grandpapa	Gravesend

He then told us the story, an incident in his own life, about a trip on the Thames and a picnic and who Granley was and so on. All of this I have forgotten, but the device and its significance has not been forgotten, partly because it was the first of this type I encountered and partly because of its peculiarity and the circumstances surrounding its first appearance.

This teacher had a great fondness for memory devices and was ingenious in their construction. He gave us a great number to help in memorizing geographical, historical and literary facts. I have now forgotten practically all of them. A slight memory of some of the mnemonic sentences remains, but these sentences are not connected with any facts. For example, there was a sentence about "little Gwendolyn" which was to help us remember the names of the Welsh mountain peaks, but this is all that I remember of the device or the mountains. After the initial novelty of the device had worn off, what I was learning was both the sentence and the facts. What started out to be an aid, became shortly a fact to be memorized. Unless somehow or other the mnemonic device is intrinsically connected with the facts themselves, it is of doubtful value.

Such an intrinsic connection is shown in this device which I remember from my school days:

> "Peccavi, I've sinned," said Lord Ellen so proud;
> Dalhousie, more modest, said, "Vovi, I've vowed."

Which tells me that Lord Ellen (and I have a faint memory that his real title was Ellensborough or something like that) acquired the province of Sindh and Dalhousie the province of Oude in India. Whatever the facts may be worth to me now, the device has preserved them for many years. Psychologically speaking the bonds required to learn the rhyme are easier than those required to learn the bare facts, and when the rhyme itself has once been learned it contains the necessary facts within it.

Devices we make up ourselves are much more effective than those learned from others or from books, hence the doubtful value of much of what is sold in systems of memory training. The self-made device is learned, so to speak, in the process of making it. The device of someone else must be learned as a new fact and is not so readily retained by the learner. An amusing example of this necessity for learning the device of another and the possibility in so doing of making mistakes can be illustrated by this experience. I found it hard in a certain city to remember the sequence of cross streets in the down-town section when riding down town on a street-car, there being no street signs at the intersections. The streets were in order named Spring, Long, Gay. So I made up the sentence "spring along gayly" and this gave me an easily remembered device. I told this to a colleague, who also claimed to have experienced the same difficulty. He thought it an excellent little memory device. A few days later I overheard him explaining it to another, but now it had become, "spring gayly along."

Memory devices, therefore, form no exception to the general rule in learning, namely, that we learn precisely what we practice. We must, therefore, be critical of accepting them, unless the device to be learned contains in itself or closely bound up with it the things we wish to remember. Mnemonic devices are not in themselves either good or bad. They do not strengthen or weaken the memory. They are just like any other information and the laws of learning function with them just as with anything else. It is quite all right to use them at times, but remember that the best are those which you make up for yourself, and the funnier, sillier, more striking they are, the more likely are they to be easily remembered. We have discussed memory devices under our topic of relevant and irrelevant reactions, because they fit in just here. If they are relevant they may be helpful; if they are irrelevant they most certainly are not helpful.

Emotion and Learning. — The great spurs to learning are feelings of satisfaction and annoyance. We must be careful to distinguish such feelings from strong emotional states. In general it is safe to say that intense emotion hinders rather than helps the learning process. Extreme fear, rage, joy, anger, hatred are not good companions for the learner. They distract the attention. They interrupt the sequence of reactions necessary for the learning of the matter in hand. Children may be so "scared" of making a mistake that they make nothing else but mistakes. The examination situation may so terrify a pupil that he makes all kinds of wrong reactions. Urging children to hurry may cause flurry and excitement. Rivalry between individuals or groups may go over into intense excitement and worry. Scolding or ridicule for failure may cause fear or anger in a child. Whenever such emotions become fairly intense, they are sure to interfere with our habitual reactions.

Emotions represent a general stirring-up of the visceral and organic parts of the individual. They represent a preparation for extreme physical exertion. They do not help in the finer reactions required for acts of skill or acts of judgment and reasoning. They disturb the poise of the learner.

The Whole and Part Methods of Learning. — Much experimental work has been done in trying to solve the question as to how much of a connected piece of work we should try to learn at once. Should we learn a poem bit by bit, line by line, or should we try to learn the whole poem as one piece. The general answer from experimental work on this problem is that the whole method is better than the part method, other things being equal. This means that wherever we have a unit of connected work to master, it is better to try to learn it as a unit, rather than to split it up into parts. Of course, the unit must not be too large. We must be able to envisage it easily in our learning period. We must be able to go over it several times at one sitting for practical school learning. What may adequately constitute a "whole" for one individual may be too much or too little for another.

The reason for the superiority of the "whole" method is really dependent upon the principle we have repeated so often in this chapter, namely, that we learn exactly what we practice. In the "whole" method we practice the poem or passage as a unit. Whereas, in the "part" method we practice the bits in isolation. After the bits have thus been learned, we have to learn them all over again as a unit in order to make them hang together. We all know how difficult it frequently is to remember the beginning of the "next" stanza in a poem. This is not only due to the difficulty inherent in the fact that frequently a new theme or thought is introduced at the beginning of a stanza, but the difficulty may also be increased by the fact that we have learned the stanzas as separate units, and the bonds between stanza and stanza are weak.

The practical exigencies of mass instruction lead us as teachers to break up our material into bits that can be assigned from one day to the next. Much of this is no doubt necessary, because we need a check-up from day to day, as to what the child is learning. For some things this is no misfortune, but for other units of work it is a drawback and we should try as far as possible to give long-time assignments of units that are "wholes." It certainly is not necessary for all the children in a class to memorize the same poems and to learn them by bits. Why not let each child choose his own poem and be made responsible for learning it by a given date, and so on with other topics.

Of course the whole method must be used with commonsense. Reading over the unit as a whole until learned is best, provided each item in the whole is equal in difficulty to all other items. Then all the items will all be learned equally well at the same time. This is never the case with material outside of the laboratory. Therefore, the "whole" procedure must be used until a general knowledge of the whole has been gained. Then we must concentrate on the difficult bits bringing them up to the whole unit. Then we must go over the whole again, and so on until learned.

The whole method is used by the young child learning the songs sung to him by his mother. He hears them over and over again and sings them over and over again with her and ultimately he has learned the whole song or nursery rhyme. He does not learn them line by line. And so he learns to read following the words of the story as his mother reads, going over the story again and again from beginning to end until one day he can say the whole story and keep the right place with the printed symbols. And the whole method is the one used by most people not in school. Folk-songs are handed down from one generation to another by the frequent repetition of the whole song. And so are stories and legends transmitted.

We don't sit down to learn these things bit by bit. The minstrels, bards, troubadours and minnesingers picked up by the "whole" method the songs, stories and legends of their profession. We learn exactly what we practice. The "whole" method is best for learning things by "wholes," and the "part" method is best for learning things by "parts." Now in most material we generally want the thing to hang together as a unit; we want to learn the whole thing, and so in most school work we should approximate the whole method as far as we can.

Most of this discussion of the whole and part method refers to memorizing things. We are not so sure about the superiority of the whole method in the case of learning complicated motor acts. It has been shown that one can make a pencil maze so complicated that it is practically impossible to learn by the whole method, whereas it can be mastered by the part method. We are, therefore, at present not certain of the respective merits of the whole or part method where motor activity is concerned.

The Length and Distribution of Learning Periods. — Short periods of practice or drill are better than long ones. Of course, the subject studied and the individuals who are studying will have to determine what is to be considered short or long in each case. We cannot say that five minutes is always better than ten minutes. Five minutes may be an adequate period for some kinds of drill, but it may be totally inadequate to get a good start in other kinds of work. The unit or "whole" which we discussed above will help to determine the length of the period. The period must be long enough to go over the work several times if it is a small unit and at least once if it is a long unit. In highly concentrated drill the periods can be made very short to be effective.

Now it follows that if we make our practice periods short, we shall have to distribute them over a greater number of days than would be necessary for longer periods, assuming that we want to devote the same total amount of time to the learning. The distribution of the periods and the length of the periods are, therefore, closely related. Arrange for short lengths of practice spaced out over a long period of time.

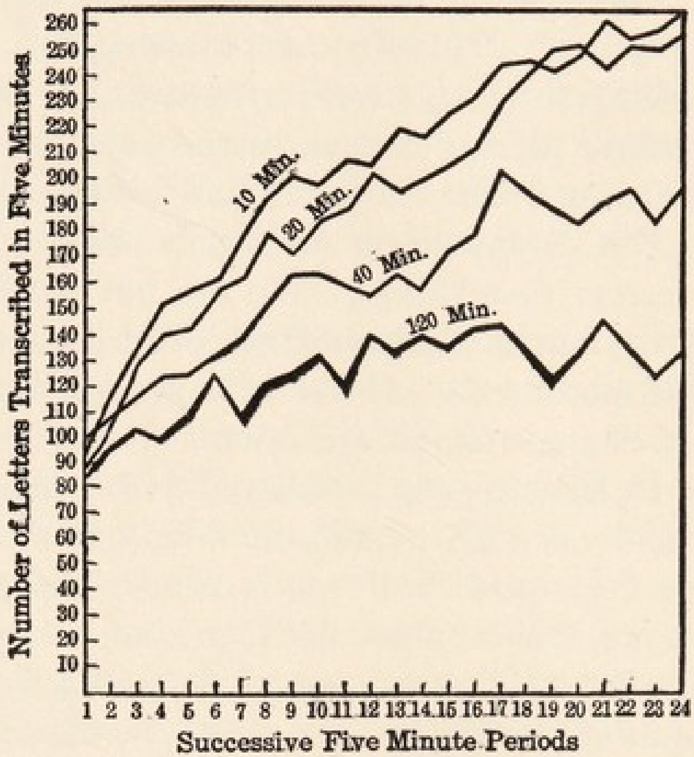


FIGURE 62. — Practice in writing letters for numbers according to a key. (From Starch's *Educational Psychology.*)

Work done in the laboratory shows this law very nicely. The curve in Figure 62 is taken from an experiment by Starch, and shows the efficiency of four groups of students in each successive five minute period in writing numbers for letters according to a key. The meaning of the curves is as follows:

10 min.	curve	=	group working		10 min.		twice a day.	
20	"	" =	"	"	20	"	once "	"
40	"	" =	"	"	40	"	every other day.	
120	"	" =	"	"	120	"	at one time.	

The two groups showing most improvement are the group which has the shortest practice period and the group which has fairly short practice periods well distributed. The twenty minute period is almost as good as the ten minute period because the distribution of the periods, one a day, is probably a little better than the distribution of the shorter period, twice a day.

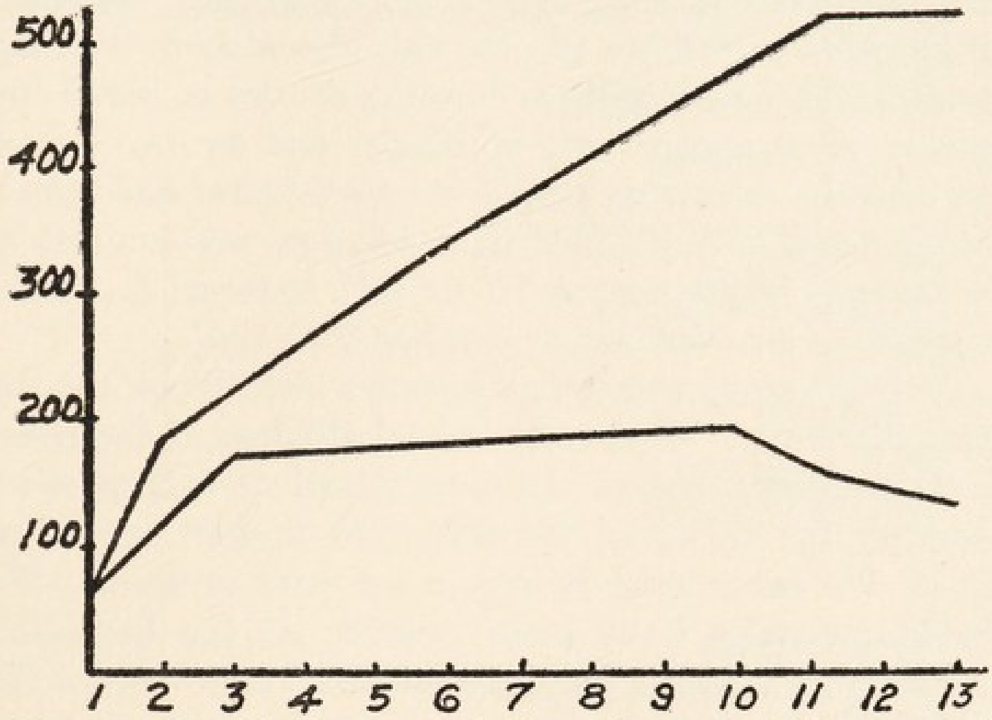


FIGURE 63. — The lower graph shows the results of fourteen half-hour practices on the same day in a letter-symbol substitution experiment. The upper graph shows the results of the same number of practices, one a day for fourteen days. (From Pyle's *The Psychology of Learning*.)

If we bunch our practice very much, if we try to cram it all in a day, we may find that after a certain period we are making no progress. We may in some functions even be getting worse. After a certain period we may lose interest, get bored and careless. Thus Pyle (6) found that 14 half-hour periods of practice in a letter-symbol substitution experiment all on the same day led to virtually no improvement after the third half-hour period and after the eleventh half-hour period his efficiency actually decreased. A very clear case of diminishing returns for his efforts. This is shown by the lower curve in Figure 63.

The upper curve in the same figure shows how improvement in the same task continues to the end of the fourteenth practice period when these periods are distributed one a day over fourteen days.

Cramming is bunching one's work in a short period of time just before it is needed. If we try to cram too much into a given period, we run the risk of confusion, forming bonds in the wrong order or forming one set of bonds too quickly after another set of bonds, and so on. If we lengthen the cramming period we get fatigued and bored. Cramming is only justified if we have no real interest in the thing to be learned, and if we wish to forget it as soon as a certain examination or test has been held.

There are many other experiments which all point in the same direction. Work with school children has verified the laboratory findings. In most school work there is no necessity for crowding the drill into a short period of time. We can spread it over a semester or term. We should naturally have more practice at the beginning and then let it taper off more or less evenly toward the end. This kind of review or re-learning from time to time is necessary in order to refreshen and strengthen old bonds. Thorndike has pointed out the desirability of considering carefully these factors in the construction of text-books in arithmetic, and he has analyzed the amount and distribution of practice with different kinds of number combinations as actually found in several arithmetic text-books. For example, he finds that in one set of arithmetic texts, the amount of practice with 5×5 gradually increases from its first appearance in grade 3 until the end of grade 6. In grade 6 the pupil would actually get a greater amount of practice in multiplying 5 by 5 than he received in any preceding grade. Figure 64 shows the distribution of practice on 5×5 according to Thorndike's analysis. From Thorndike's many diagrams we may justly conclude that little thought has been given to this problem of the amount and distribution of practice in most arithmetic texts. Now that this has been pointed out, future books will probably be better planned.

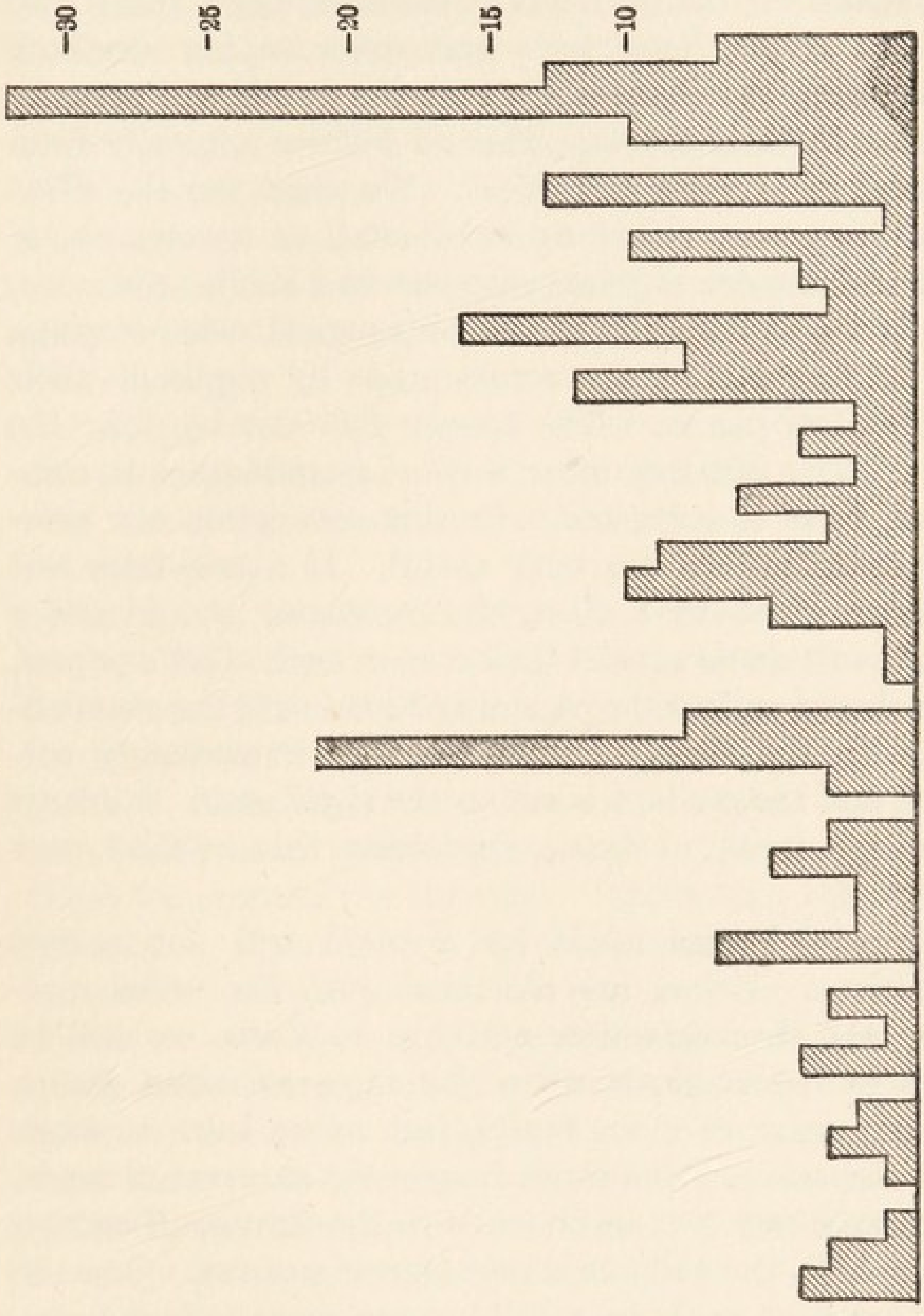


FIGURE 64. — Distribution of practice with 5×5 in the first two books of the three book Arithmetic text E. (From Thorndike's *Psychology of Arithmetic.*)

Knowledge of Improvement. — Keep the pupil informed of just what improvement he is making. If he knows he is succeeding, that will make further success easier. It has been shown in many laboratory experiments that subjects do not improve so fast if they are not told their scores from time to time. If you give a test and do not tell your pupils their scores, you will find no enthusiasm for the next test of the same type, and if you keep on giving these tests and never saying anything about them, the chances are that the pupils will actually decrease in their scores. This all follows naturally from our laws of exercise and effect. We must feel the effect of our response, either as satisfaction or annoyance, in order to influence our next response to a similar situation. This is the explanation of the pedagogical value of going over and explaining the errors made by pupils on their examination papers. The sooner this can be done the better. Immediately after a short examination is completed, have it corrected. In this connection the new-type examinations are very useful. If a true-false test has been given to a class, the instructor should either have the students correct their own or each other's papers, or he should collect the papers and read over the examination with the class. Errors are thus immediately corrected and the student is set on the right path. Students of college level, at least, appreciate this method very much.

Errors. — Errors made by a pupil will not correct themselves. Errors are reactions, just like other reactions, and if satisfaction attaches to them, we will be likely to repeat them when the same situation recurs. We may learn to make errors, just as we learn to make right responses. The more frequently an error is made, the more likely is it to be made in the future. Teaching should help the child to avoid learning errors. Once an error has been made by a child, there seems to be a possibility of its recurring. Only slowly is it stamped out. Myers (5) recorded over a period of time the following responses made by a child to the combination 4+6:—8, 8, 10, 8, 10, 10, 8, 10, 10, 10, 10, 8, 10, 10, 10, 10, 10, 10, 10, 8, 10, 10, 10. In this series we see the persistence of the wrong response 8, for a long time. We must attach satisfaction to 10 and annoyance to 8 in order to encourage the one and discourage the other.

Speed and Accuracy. — Accuracy in most work is desirable. In some work, such as arithmetic, accuracy is all important. It is no use teaching a child to be 50 or 75 per cent accurate in his arithmetical responses. We must strive for 100 per cent accuracy or as near to that as we can attain, otherwise there is little value in the arithmetic we teach him. In other subjects 100 per cent accuracy may not be as important. We can profit a great deal from what we read, even if our comprehension is not 100 per cent accurate. We do not need 100 per cent accuracy in writing, whatever that may mean. And so with other subjects.

But speed, too, is valuable. The slow worker falls behind in the competition of life. He is ineffective because he is slow. There is so much to learn, there are so many modifications to be made, that speed in learning to make them is desirable. The old proverb says, "Slow but sure," and many people believe this. They have further taken for granted the converse, "Quick and inaccurate." Fortunately, psychology has demonstrated that both of the statements are wrong. To be sure there are all types of individuals, slow and accurate, slow and inaccurate, quick and accurate, quick and inaccurate. In general, however, those who are quick tend to be more accurate than those who are slow, and those who learn quickly retain more than those who learn slowly. Speed and accuracy go together more frequently than slowness and accuracy.

The quick reaction is the well-learned habitual reaction. The slow reaction is the partially learned, doubtful reaction, and is more likely to be inaccurate. Compare the quick accurate reactions of the practiced adult reader with the slow inaccurate reactions of the first-grade beginner. Slowness is generally indicative of hesitancy, uncertainty. It is the symptom of the learner, whose habits are not yet well-formed. Of course, we may learn to react slowly, just as we may learn to react quickly. Other things being equal, therefore, we should try from the beginning to form quick reactions, rather than slow ones. Once we have learned the right response, we should speed it up before we get into the habit of reacting slowly. Slow reactions are frequently indicative of less desirable methods of response than quick reactions. The child who reacts slowly in adding, for example, probably does so because he is counting on his fingers, or counting to himself. He is employing a less desirable form of reaction than the direct response between two numbers and their sum. And this less desirable response not only takes longer but is subject to more possibility of error.

Many workers have shown that in most school subjects, children can easily be taught to increase the speed of their reactions and this without any loss in accuracy. Some have found increased accuracy resulting from the speeding up process, very probably because the children have substituted more desirable for less desirable methods of response.

Pupils in the grades can be taught to increase their speed of reading about fifty per cent without disturbing their ability to comprehend. This is an enormous gain in speed without any counterbalancing disadvantage.

Figure 65 shows different speeds and qualities in measuring the work in silent reading of 1831 pupils. The numbers in the circles indicate percentages. The smallest percentages are for the "slow but good" and the "rapid but poor." High rate and good quality are generally related.

Age and Learning. — As we grow older we feel that it becomes more and more difficult to learn new things.

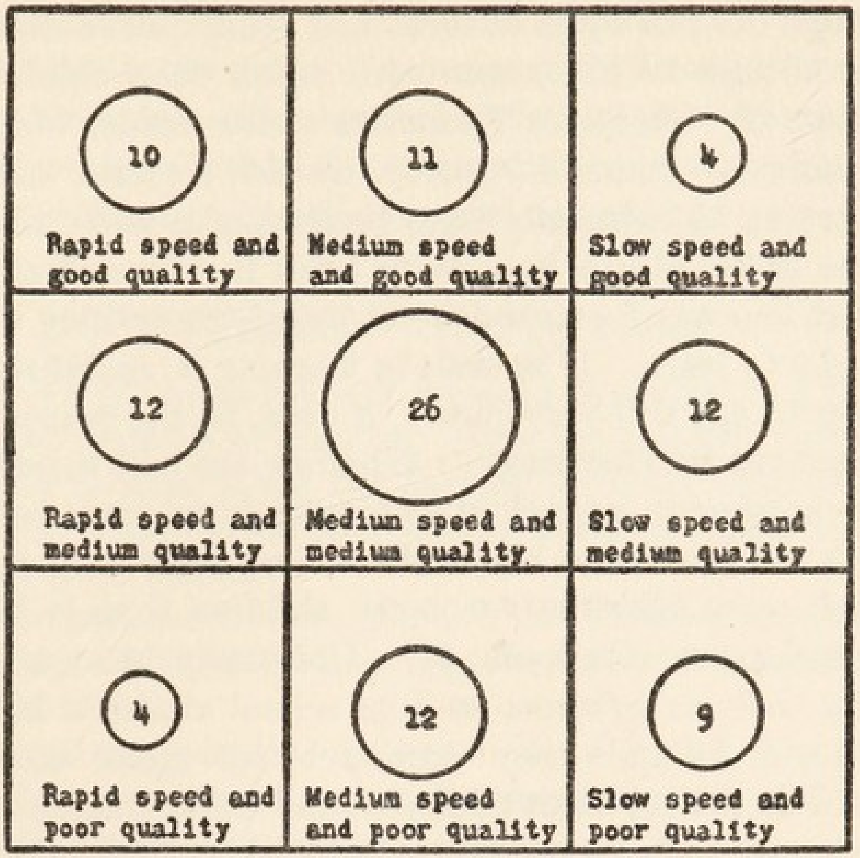


FIGURE 65. — Per cent of 1,831 Cleveland pupils found in each of nine speed and quality groups in silent reading. (From Judd's *Measuring the Work of the Public Schools.*)

"You cannot teach an old dog new tricks." This is only partially true. What happens to most adults is that they do not want to learn; or better, they know more definitely than the child does what they do and what they do not want to learn. The child is more docile. Thorndike's (7) studies of adult learning show that ability to learn increases up to late adolescence. There is no indication of a decrease in this ability up to age 25. From age 22 to age 42 the decline is very slow, only about one half of one to one per cent a year. This slight decrease is about the same for inferior and superior intellects, although they differ greatly as to the difficulty of the tasks they can master. It is only when we reach the age of 55 or thereabouts that ability to learn seems to be definitely on the wane. Childhood is not the easiest period in which to learn a language. Given equal interest and equal native capacity adults of ages 20 to 40 can learn better than children of ages 8 to 12. To quote Thorndike's own words, "In general, nobody under forty-five should restrain himself from trying to learn anything because of a belief or fear that he is too old to be able to learn it. Nor should he use that fear as an excuse for not learning anything which he ought to learn. If he fails in learning it, inability due directly to age will very rarely, if ever, be the reason."

Incentives for Learning. — What are the best incentives for the teacher to employ in making children learn better? As between praise and blame there is no doubt that praise is much more effective for most children than is blame or punishment (Hurlock, 4). University students reporting their experiences as high school students believe very decidedly that they worked better under such incentives as praise, encouragement and commendation and worse under such incentives as censure, ridicule, threats, sarcasm and punishment. Censure and punishment may at times be useful and necessary, but in general there is no doubt that the teacher should rely on encouragement and praise.

Mental Imagery. — In older text-books of educational psychology a great deal used to be said about mental imagery and its effect on the learning process. It used to be thought that each individual was dominated by a special type of imagery, and such expressions as "eye-minded," "ear-minded," and "motor-minded" came into use. As a matter of fact, we now know that pure types rarely exist. All of us use all types of imagery. Children are not sharply divided into visuals, audiles, motiles and so on. It is best for the teacher to disregard these distinctions, and it is almost impossible anyway to determine whether a child uses predominantly one type of imagery rather than another. Attempts at classifying pupils into image types should be abandoned. Material should be presented in the most effective ways in order to call forth the kinds of reactions which will be required later on by the child. Only in cases of special inability to learn need we try other modes of presentation than those usually found to be effective and economical.

Summary — How to Study

All of this chapter has been devoted to a discussion of the most effective methods of learning. From the point of view of the student or pupil doing the learning we could summarize the material of the chapter into rules for studying. These might be somewhat as follows:

1. Learn to react in the way you will be called upon later to react. This is the most important and fundamental of all the rules.
2. Don't learn irrelevant details. Avoid irrelevant responses.
3. Don't set yourself an aim impossible of attainment. Don't use a model impossible to achieve.
4. Don't sit and wait for inspiration. React, and inspiration may come later.
5. Use only such memory devices as are a real help. Don't waste time learning the devices. The best are such as you make up for yourself.
6. Intense emotion of any kind hinders learning. Try to work in a calm manner. Don't fuss and worry.
7. Learn whole units as wholes, so far as is feasible. Don't chop up natural units into unnatural bits.
8. Devote a reasonable amount of time to any particular task. Let it be short rather than too long. Short periods are better than long.
9. Distribute the periods of study over as many days as possible. Ten minutes each day for six days is better than sixty minutes all on one day.
10. Cramming is bunching your work all together just before it is needed, and so it is uneconomical as we saw in rule 9. But if you don't want to retain it and have to learn it, nevertheless, cramming is the quickest way to get done with it.

11. Keep a record of improvement. Watch your score. It will stimulate you to do better.
12. Avoid error from the start. Do not be complaisant about errors. Making errors means you are forming a habit of error.
13. Speed and accuracy go together. Don't dawdle. Keep moving. Most likely you can increase your speed and increase your accuracy at the same time.
14. People are never too old to learn. Older people can learn about as rapidly as younger people.
15. The best incentive for learning is praise rather than censure. In your teaching use ridicule and sarcasm sparingly.
16. Don't bother about your mental imagery. You probably use all kinds. Forget it.
17. Don't try to learn too many things at the same time. Wait until you are well along in one thing before adding a new subject.
18. After working intensely at one topic, leave it alone for some time. Give it a chance to form connections with the rest of your knowledge and thus become a real addition.
19. If you have an essay to write, jot down at once your ideas about it. Keep thinking about it at odd moments and adding to your notes. You will be surprised how much you can do in preparation before you actually sit down to write it.

Review

True-False Statements

As a review mark the following statements true or false:

1. Teaching young children to trace letters will speed up the initial stages of writing instruction.
2. Strong emotion stimulates mental activity and leads to greater achievement.
3. Distributed practice periods usually yield better returns for time spent in learning than does concentrated practice.
4. The best way to learn a thing is exactly in the form it is to be used later.
5. An undesirable response may be eliminated by setting up in its place a more satisfying one.
6. The so-called "natural methods" of learning, which a pupil discovers for himself, may not be the best for him.
7. Intense emotional excitement is not helpful to learning.
8. To learn by the "part" method is always more effective than to learn by taking large "wholes."
9. It is psychologically wrong to use memory devices as an aid to memory.
10. Learning should be thought of as reaction rather than absorption.
11. It is important to know what habits should be formed and also the order in which they should be formed.
12. It is well to explain the theory of long division before giving your pupils examples to practice.
13. Children who work rapidly usually make the greatest number of errors.
14. Learning by doing is based on a false psychology.
15. Children can learn poetry better and more rapidly at age 9 than at age 16.
16. It is good to have frequent reviews of a subject at short intervals throughout the semester.
17. Study your French lesson "out loud" if you want to learn to speak French.
18. The laboratory method of teaching is always superior to the lecture method.

19. A self-made memory device generally aids retention less than one of the standard devices.
20. Cramming is psychologically and morally wrong.
21. The teacher by not giving out the results of achievement tests sustains the interest of his pupils.
22. After forty an individual is too old to begin learning a foreign language.
23. Most pupils can be divided into two classes, strongly "ear-minded" or strongly "eye-minded."
24. Ridicule generally acts as a decided spur to further improvement.
25. Check the best response:
    Speed and accuracy in the performance of a task:
    (a) seem to have no connection with one another.
    (b) generally go together.
    (c) never go together.
26. Check the best response:
    Learning is a process of:
    (a) passive absorption.
    (b) eliminating original tendencies to react.
    (c) forming new connections.
    (d) storing up knowledge.

Advised Readings

Whipple, G. M.: *How to Study Effectively*, Second Edition, Public School Publishing Co., Bloomington, Ill., 1927.

Kitson, H. D.: *How to Use Your Mind*, Philadelphia, 1916.

Kornhauser, A. W.: *How to Study*, University of Chicago Press, 1924.

Thorndike, E. L.: *The Psychology of Arithmetic*, New York, 1922.

References

1. Gates, A. I.: *The Improvement of Reading*, New York, 1927.
2. Gates, A. I.: *New Methods in Primary Reading*, Teachers College, Columbia University, New York, 1928.

3. Hertzberg, O. W.: *A Comparative Study of Different Methods Used in Teaching Beginners to Write*, T. C. Contributions to Education, No. 214, Teachers College, Columbia University, 1926.
4. Hurlock, E. B.: "The Value of Praise and Reproof as Incentives for Children," *Archives of Psychology*, No. 71, July, 1924, Pp. 78.
5. Myers, G. C.: *The Prevention and Correction of Errors in Arithmetic*, Plymouth Press, Chicago, 1925.
6. Pyle, W. H.: "Economical Learning," *Journal of Educational Psychology*, 1913, 4, pp. 148–158.
7. Thorndike, E. L. et al.: *Adult Learning*, New York, 1928.

CHAPTER XI

THE PERMANENCE OF MODIFICATION

Disuse Weakens. — Use strengthens. Disuse weakens. As soon as we stop practicing or exercising, we begin to forget. This is the opposite or converse of the Law of Use. Whatever habits we have formed, whatever things we have learned, begin to weaken as soon as we stop making use of them. Some of our habits and skills have had so much exercise, and some of our knowledge has been so thoroughly learned that, even after long periods of disuse, there is little forgetting. With things not so well learned, however, forgetting comes quickly and is easily measurable.

Curves of Forgetting. — Just as our curves of learning pictured the growth or improvement in a function because of exercise, so we may draw curves showing the decay or weakening of a function through lack of exercise. The classical curve of forgetting is that of Ebbing-

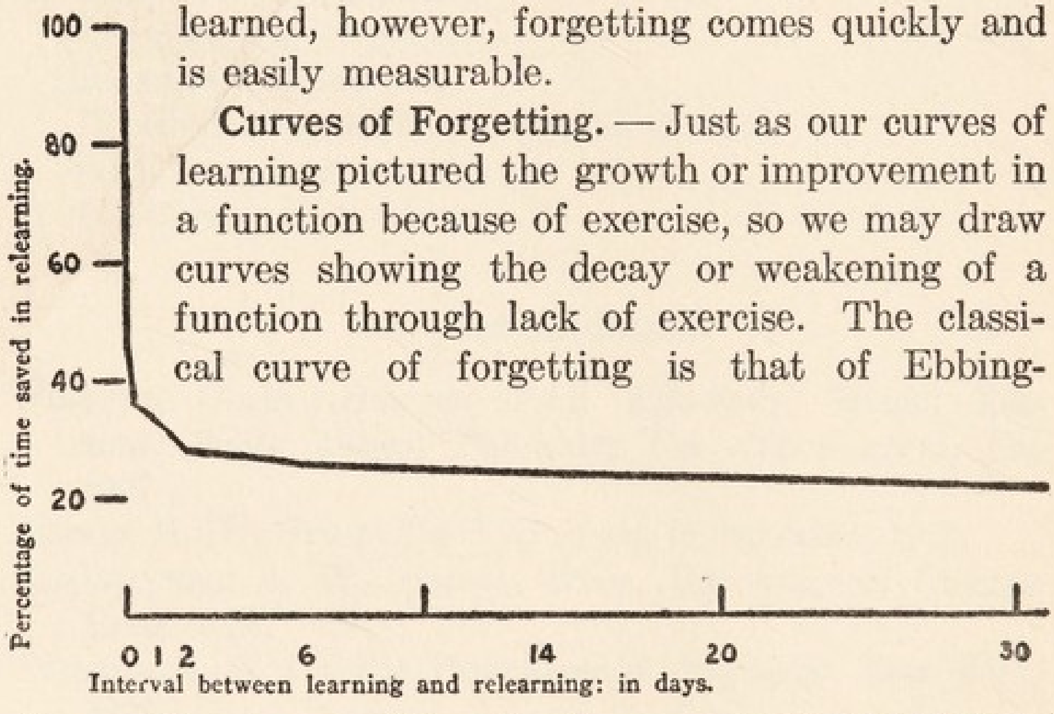


FIGURE 66. — A curve of forgetting for non-sense series by Ebbinghaus (after Thorndike).

haus, shown in Figure 66. Ebbinghaus learned a series of non-sense syllables until he could repeat it once correctly, and then, after a definite interval of time, relearned the same series, measuring how long it took him to relearn. Then he learned another similar series and after a different time interval, relearned that series. And so on for many different time intervals. In this way he arrived at the curve in Figure 66, which shows that the longer the interval before relearning, the smaller the percentage of time saved in relearning, or the greater the effect of forgetting. We see, therefore, that with this kind of learning at least, forgetting is very rapid at first and gradually becomes less and less rapid. Even within an hour after learning the words, more than half had been forgotten.

Other investigators have not found forgetting to proceed quite so rapidly as Ebbinghaus did. Another curve (4) for forgetting is given in Figure 67. This shows the percentage retained after intervals of 30, 60, 90 and 120 days. The students who took the experiment observed a card with a number of objects and photographs on it and then wrote out what they had seen. The curve is not nearly so steep as the Ebbinghaus curve.

We cannot take Ebbinghaus' results and frame a general law. There are evidently all kinds of forgetting curves, just as we found there were all kinds of learning curves. The more thoroughly the thing has been exercised or learned the less steep will be the forgetting curve. Ebbinghaus' curve represents the loss for things just learned to the bare point of retention. Not all that we learn is learned in that way. Frequently we over-learn quite a bit and in some complex functions there are parts which we over-learn tremendously, so that they are retained for very long periods of time.

Over-learning is very obvious in complex acts of skill, such as typewriting, swimming, driving an automobile and so forth. In such complex functions many minor habits are tremendously over-learned, so that they may not be forgotten after years of disuse. The total complex

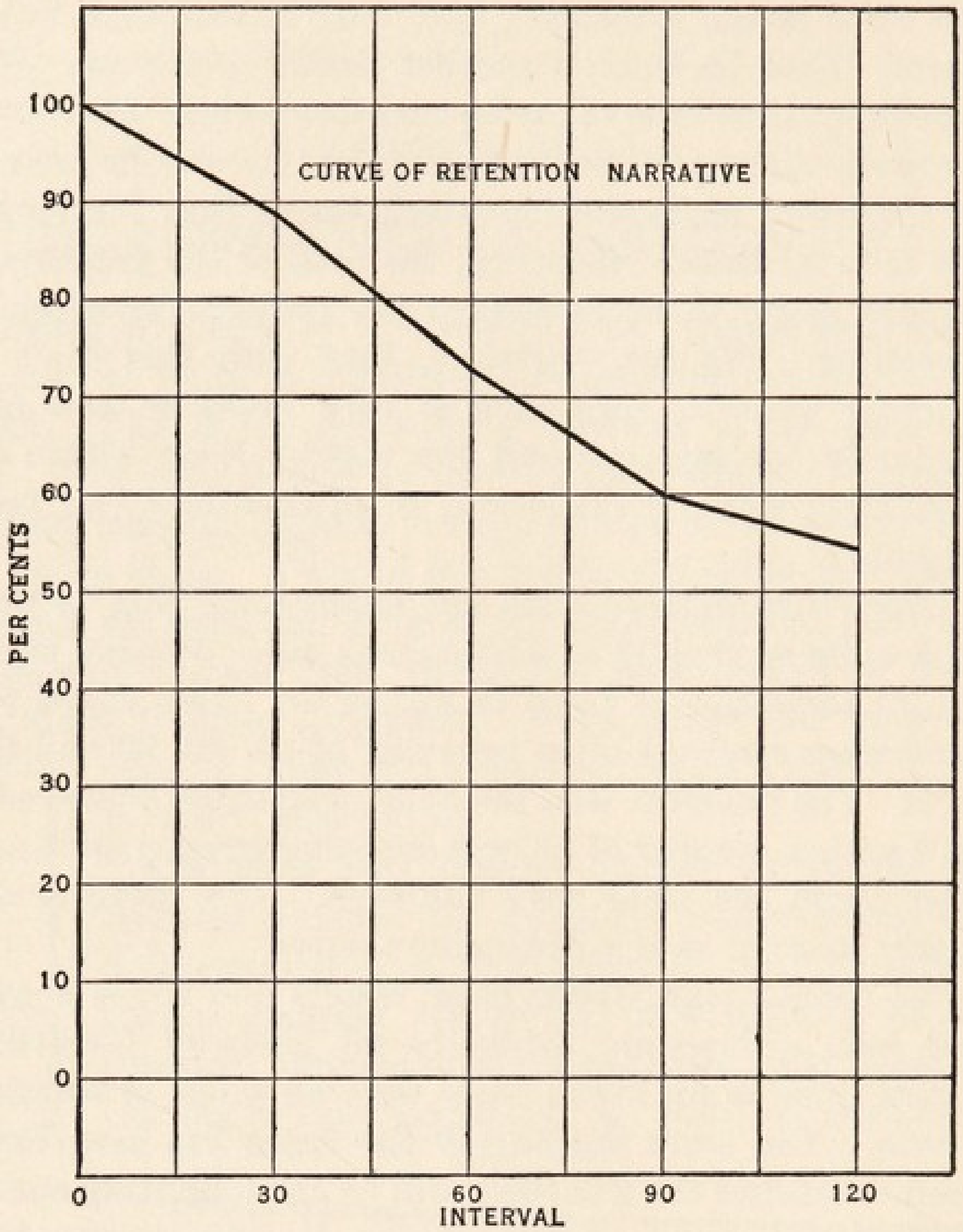


FIGURE 67. — A curve of forgetting for meaningful material. (From McGeoch and Whitely, *Journal of Educational Psychology*, 1926.)

function does not drop out of existence like a series of non-sense syllables. Much of it is still left. Enough of it is left, so that we do not say, "I have forgotten how to typewrite, or how to swim," but rather, "I'm very rusty in my typing or swimming." Enough of the connections will still function, so that after some exercise we may be able to bring ourselves up to the level of efficiency attained before and in some cases to exceed that level. Book (1) for example, found that after a period of seventeen months without practice in typing he had lost only a little of his skill, and that after ten daily practices of 10 minutes each, he was able to exceed his best previous record. Here the loss from disuse was very rapidly regained.

This ability to reinstate rapidly things partially forgotten, which had been previously well exercised, is familiar to all of us. Much of our school knowledge may be forgotten, but how quickly it comes back to us if we go over it again. Many a father or mother is surprised at the amount of Latin or algebra he or she knows when going over it again after many years of disuse with a son or daughter. Much of it comes back quickly and the parts that were frequently exercised come back the most readily of all. The Mother Goose rhymes you learned in your childhood may not come back readily just now, but if you give them a little encouragement they can soon be brought back to their former strength, as many a young mother can testify.

The significance of all this is that we should not worry too much about forgetting. Many reactions now forgotten can easily be revived when needed. If we wish to keep a reaction continually ready to function, after once being thoroughly learned, just a little exercise now and again, lest we forget. No improvement, habit or knowledge is absolutely permanent, and yet on the other hand neither is it ever absolutely forgotten. It has left its trace and will be more easily revived than it was originally learned.

We cannot say, however, that all practice always leads to permanent modification or change. The effects of some practice seem to be very evanescent. In an investigation (Gates and Taylor, 2) with a group of school children, 78 days of practice in memorizing digits was given.

During this period of use the average score rose from 4.36 to 6.36. Then followed a period of disuse or no practice for about five months. When tested after this period, the average score dropped back to 4.71, or almost as low as at the beginning of the whole experiment, before any practice at all. A little later the practice was resumed and in 22 days the score rose from 4.71 to 5.73. In the first period of practice it took 44 days to arrive at this score of 5.7, but this time it only took 22 days. At a first glance this would seem as if the first long practice had made the relearning easier. This was not the case, however, for a control group which had no previous practice raised their score from 4.83 to 5.92 within 22 days. In other words, they did just as well as the experimental group without the benefit of the long period of 78 days of practice. In some things, therefore, skills or techniques may be learned and readily forgotten, so that in beginning again we have practically to relearn them all over again.

Supposed Contradictions to the Law of Disuse. — It has been suggested that there are some exceptions to the general Law of Disuse. Thus it has been said that "we learn to skate in summer and swim in winter" meaning thereby that a period of disuse may lead to improvement in a function. There is little experimental evidence for anything like this. It may be true that there have been cases where, after a period of disuse, an individual did better at a task than he did just at the point of stopping. It will generally be found, however, that this stopping point was by no means the highest point reached by him in his improvement. He had previously reached higher points and his ability after disuse had not exceeded his highest level of performance. When the golfer makes a good score on his round after a period of lack of exercise, he may compare it with his last score at the end of a season's play, when he was fatigued or "stale" from too much play. If he compared it with his best score during the past season, he would find that disuse has weakened the bonds somewhat.

If there are any authentic cases of improvement after disuse, they must be explained on the theory that lack of exercise caused the diminution of undesirable bonds which were impeding improvement. When these had been weakened by disuse, the desirable bonds were able to function better than they ever did before.

School Work. — Forgetting of knowledge and skill learned in school by the child goes on steadily, as soon as lack of exercise takes place. Hence the necessity for frequent review. Hence the necessity for the distribution of practice over long intervals. During the long vacation from June to September forgetting takes place. The actual loss in certain school subjects has been measured. Every teacher knows the necessity for some relearning in September. But if the previous learning has been efficient and economical, a little relearning will bring it all back.

One investigator (5) gave standard tests at the end of the school year in June and then again at the beginning of the next school year in September. He found a slight gain on intelligence and reading tests but a slight loss on arithmetic tests. Evidently the kind of reaction measured by the intelligence and reading tests is being exercised during the vacation and so these abilities continue to increase as the child develops. Very little of the formal arithmetic practiced in school is made use of during the long vacation, and therefore it tends to diminish by disuse.

The Effect of Lectures. — How much does a college student retain as a result of listening to a lecture and how long is it retained? Careful measurement (3) of many groups gives the curve shown in Figure 68. Immediately after the lecture 62 per cent can be recalled. Three or four days later about 50 per cent is retained. After one week the percentage drops to 37, after two weeks to 30 and after eight weeks to 23. Like the Ebbinghaus curve we have here a relatively sharp drop at the beginning and then a more and more gradual decrease.

Speed of Learning and Retention. — Does the rapid learner forget quickly? Does the slow learner retain most? Evidently not. The indications are that those who learn quickly in general retain most, and probably also retain it for a longer period.

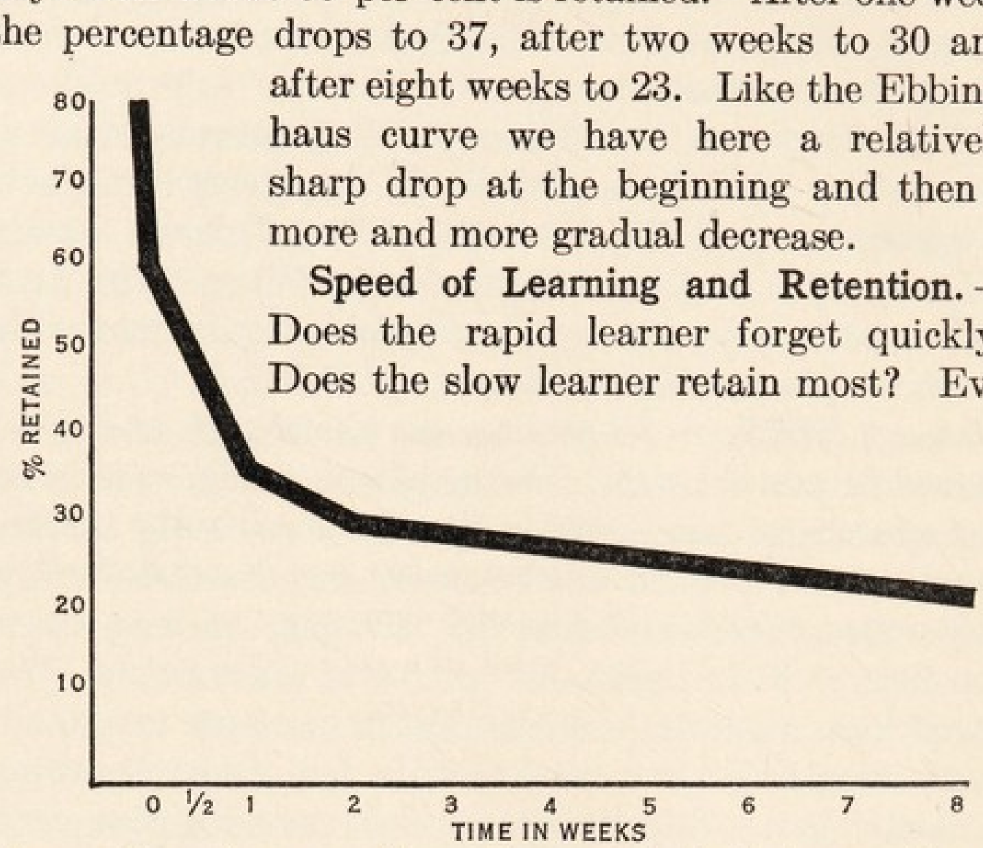


Figure 68. — A curve of forgetting material in lectures. (From Jones' *Experimental Studies of College Teaching.*)

Over-Learning. — This is a phrase used to express the fact that we may learn something past the point of just being able to retain or accomplish. The longer we wish to remember something, or the more permanent we wish to make a modification, the more must we over-learn. We have had a great deal of practice with the multiplication tables. We have over-learned them tremendously. If we were not to make use of $3 \times 3 = 9$ for many years, we would still be able to respond with 9 at the sight of 3×3, because we have over-learned this very greatly. In cramming we do not over-learn. We learn just about to the point of bare retention, and hence things we have learned by cramming are soon forgotten.

Summary

1. Disuse weakens. Forgetting begins as soon as we stop practicing.
2. The rate at which we forget things varies greatly. Some modifications are more permanent than others. There are all types of forgetting curves.
3. Permanence is greatly influenced by the amount of over-learning. Some facts and some skills are very permanent, because they have been so thoroughly over-learned.
4. What has been thoroughly learned and then forgotten or partially forgotten, can generally be very rapidly reinstated.
5. During the summer vacation school children tend to forget much of that which they do not make use of during this period but they continue to gain in such subjects as reading.
6. The memory of things learned in a college lecture decreases steadily for about two weeks. Nevertheless, after eight weeks about 23 per cent is still retained.
7. The rapid learner retains most and also probably retains it for a longer period.

Review

True-False Statements

As a review mark the following statements true or false:

1. During a period of disuse a function remains stationary.
2. The slow learner retains what he gets better than the quick learner.
3. The Ebbinghaus curve is the typical forgetting curve.
4. If we wish to retain for a long period of time, we must over-learn considerably.
5. In complex acts of skill, many parts are generally considerably over-learned.
6. The more a thing is over-learned, the more readily can it be reinstated once it has been forgotten.
7. There is some evidence to make us believe that some skills go on developing or increasing during a long rest period.
8. In September after the summer vacation children are just about as good in arithmetic as they were in the previous June.
9. "Easy come, easy go," meaning, that those who learn quickly, forget quickly, has been shown to be true by psychological measurements.

Advised Readings

Thorndike, E. L.: *Educational Psychology: Briefer Course*, Teachers College, Columbia University, New York, 1914, Chapter 17.

References

1. Book, W. F.: *The Psychology of Skill*, New York, 1925.
2. Gates, A. I. and Taylor, G. A.: "An Experimental Study of the Nature of Improvement Resulting from Practice in a Mental Function," *Journal of Educational Psychology*, 1925, 16, 583–592.
3. Jones, H. E.: "Experimental Studies of College Teaching," *Archives of Psychology*, No. 68, 1923.
4. McGeogh, J. A., and Whitley, P. L.: "The Recall of Observed Material," *Journal of Educational Psychology*, 1926, 17, 419–425.
5. Morrison, J. C.: "The Effect of the Summer Vacation on Learning," Ohio State University, *Educational Research Bulletin*, Oct. 1, 1924, 245–249.

CHAPTER XII

THE TRANSFER OF TRAINING

Reciprocal Modification. — So far in all our discussion of the modification of original nature, we have taken up one modification at a time. All the learning curves in Chapter IX dealt with the modifications of one function, with the learning of a definite task, with the formation of a specific habit. The learning of anything may last for many days or months or years, as when one learns golf or arithmetic or good manners. And in between the practice of any of the specific functions we are learning other things. We do not practice one thing until we are perfect, and then begin another. In every-day life we are practicing countless things at the same time. And so, in this chapter, our problem is to discuss what effect modifications of original nature have upon each other, the reciprocal action of one modification upon the other. If modifications in a boy's reactions involved in golf, arithmetic and table manners are taking place during the same period of time, our problem is to discover whether the modifications taking place in one of these functions influence the modifications involved in any of the other functions, and to what extent.

If we turn to specific school illustrations, the importance of the problem will immediately become obvious. Does the learning of French help the student in the learning of another foreign language? Does Latin or grammar help a student to write better English composition? Does improvement in accuracy in arithmetic lead also to improvement in accuracy in Latin? If we train a child to be accurate in one thing, say Latin translation, does it mean that he will automatically become accurate in all other things in life? If we teach a child about honesty, will he act honestly? If we train him to be honest at home, will he be honest in school, and if so, will he be honest later on in business or in his professional career? Or in other words, what influence has any one modification upon any other?

General Assumptions in the World at Large. — Let us see if a cursory glance over many of life's activities will give us a clue to what general experience has taught the human race. Do we assume transfer or not? In many cases it would seem that we do. The good teacher is promoted to be a principal. The duties of a teacher and those of a principal are widely different, and yet the assumption is that efficiency in the first post is a good indication of efficiency in the second. So the capable university professor, whether he excels as teacher or research worker, is promoted to the administrative office of dean or president. The good salesman is made sales manager. The good lawyer, business man or politician is elected mayor, or governor or congressman. Indeed the whole theory of electing men to various offices, as in our democratic state, assumes a great amount of transfer. Because A has been a good business man and has shown ability in looking after his own affairs, therefore, he will be able to look after other people's affairs equally well. B is a good physician, and, therefore, should prove an excellent public health officer. And so we might go on with numerous illustrations of how the world at large believes and acts upon the assumption that demonstrated ability or efficiency in some particular narrow sphere is a good index of future efficiency in another sphere, even though the latter may be rather far removed from the former. Obviously the world believes to a certain extent in the good all-around man, and in his ability to function well in many capacities.

But the world at large does not carry this to extremes. To demonstrate one's efficiency in one job is not an *open sesame* to every other job. The good lawyer is not appointed public health officer. The engineer is not expected to function as minister or judge. Because one is good in any of these special lines, one is not expected to be expert in any or every other special line. The increasing number of government, state and municipal positions put under a civil service control, is evidence of the lack of belief in transfer. The city manager movement may be interpreted as a doubt as to whether the generally good all-around man has sufficient expert knowledge to manage the affairs of a city. School superintendents to-day are chosen with more care as to their particular qualifications. Not any good man will do. Obviously many positions require a special expertness for success, and expertness in any other line is no guarantee of success in general. Obviously, therefore, every-day life affords plenty of examples where transfer is not assumed.

There are other conflicting experiences of mankind that make one hesitate about assuming too much transfer. We all know that individuals may be courteous, kind and generous in company or in business relations, and at the same time be rude, cruel and selfish at home. An unscrupulous business man may be the very essence of honesty in his relations with his friends, his club, his church. The kind doting father may be an employer of children under sweatshop conditions. So kindness, honesty, generosity are not traits that transfer from one set of circumstances to all others.

It might be more correct to restate all this by saying that we find in the world around us belief in a vast amount of transfer from some positions to others, belief in moderate amounts of transfer in other cases down to a belief in no transfer in many other cases. Belief in all degrees or amounts of transfer could probably be found and such belief is based upon the general experience of mankind. Undoubtedly many such assumptions are erroneous or distorted. We probably assume more transfer in many cases than really exists.

But the general picture that this survey of everyday life has shown is a true one, namely that different amounts of transfer from one function to another exist. The scientific question is not whether there is or is not transfer. We have nothing to do with the question, "Do you believe in the transfer of training?" The only real question for psychology is, "What amount of transfer takes place from this function to that function under these specific conditions?" And the answer may be in terms of positive, zero or negative amounts.

Historical Importance of Problem in Education. — The general problem we are discussing has figured large in the history of education. There it is known under the name of the doctrine of Formal Discipline. The mind needs training or disciplining. Certain subjects are better disciplines than others, hence they should be preferred in the training of children.

This disciplinary conception of education comes into prominence in the seventeenth century, although the general theory had appeared at various times previously. It springs up to defend some well-established subject in the curriculum whose usefulness is being attacked. Before the Reformation, Latin was the most useful, utilitarian, practical subject in the curriculum for the majority of children in school, because Latin was the language of religion, of philosophy and of practically all culture. With the rise of the several vernaculars with literatures of their own, with the increasing spread of literacy beyond the ranks of the clergy and the scholar, Latin became of less practical use for the average boy in school. Hence its usefulness was questioned and its answer was that it possessed more disciplinary value than any other study. It trained the mind. Even if the content learned were useless, the training was priceless.

The doctrine of formal discipline fitted in well with the faculty psychology of an earlier period. Reasoning, memory, imagination, perhaps accuracy, are faculties or powers of the mind. What better training for these faculties could there be than Latin translation and Latin syntax? When once trained these faculties can cope equally well with any other content.

Formal discipline assumes a very great amount of transfer. The general argument of formal discipline is always employed in one way or another, whenever a well-entrenched subject is attacked by a new and more utilitarian subject clamoring for admission into the curriculum. Latin was so well entrenched and has been so well taught and systematized for so many years that it has borne the brunt of the attack. Latin has made claims to a greater amount of transfer than most of the other subjects. But really every subject makes use of the doctrine of formal discipline to justify itself in the face of an attack from a more utilitarian subject.

Science came into the curriculum as a utilitarian subject in the face of much opposition. It split up into several studies which are now being attacked by the more utilitarian General Science. Modern Languages were introduced because of their utilitarian value, but they fall back to-day upon their general cultural value or their transfer value in order to justify their place. Industrial training is attacking manual training which justifies itself on formal disciplinary lines.

As soon as a subject is introduced into the curriculum of any school or college, it begins to be organized and "formalized" by the teacher. To the teacher this subject is necessarily important and it takes the center of the stage. The teacher believes in the intrinsic value of the subject itself and he sees also the multifarious connections with other subjects. Hence the subject is not only intrinsically valuable but has transfer value to other subjects and to life in general. No one can be a good teacher of a subject without believing in the value of his subject far beyond the mere content he is teaching. I believe a course in educational psychology is not only valuable for the immediate facts it teaches, but also because the general principles developed in such a course may help the teacher in many a concrete classroom situation. I believe the point of view developed spreads over or transfers to many other situations in life. At the same time I know that I am far too optimistic in my belief in the transfer value of educational psychology. I feel sure that actual measurements of the transfer effects would be very disappointing to me as a teacher. And yet I go on hoping for these transfer values because it is only in so far as I do hope for them and believe in them that some transfer is attained.

Mental Discipline in English Public Schools. — This doctrine of formal discipline has had and still has a tremendous hold on education. It still dominates much of our education. In "A School Master's Diary" (2), published in 1918, a teacher in one of the English Public Schools records his futile attempts at innovations and bursts forth in the following paragraph:

"Subjects are taught just in so far as they are distasteful; the fact that one can work hard at anything just because it is interesting is regarded as impossible. If one begins to argue, you are countered by the shibboleth of 'mental discipline,' which is supposed to be the final word on any topic of controversy. If grammar grind provides a mental discipline, grammar grind must therefore be invaluable, quite apart from its utilitarian aspect. Consequently boys are taught many things which serve no useful purpose and lead nowhere, simply because it is good for them to have to perform arduous, pointless tasks without asking the 'why' of them."

Wide Claims for Transfer Still Current. — We must not think, however, that this belief in wide transfer is merely a remnant of a past or alien civilization. It continually crops up in our own country in many forms and in sweeping dogmatic statements. Thus we read in 1928, "A knowledge of a foreign language contributes in an unusual degree to the making of internationally-minded, broad-thinking, intellectually resourceful and contented citizens." Would it were so, indeed! And again in another connection we read, "My claim would be that a student trained in mathematics to distinguish necessary from sufficient conditions would normally inquire *in any situation* whether a thing known to be sufficient were or were not a necessity." And again about Greek: "It forces students to think closely about the meaning of words. In ordinary life people do not think about words. . . . Many people live mentally in a sort of fog most of the time. Greek forces one to express oneself accurately." These are just a few of the many modern examples of claims for wide transfer that are still current.

Lack of Transfer Noticed. — Although belief in wide transfer was general from the Reformation down to the end of the nineteenth century, there arose every now and again some doubt in the minds of educators. In Boston in 1845 a committee made a survey of the educational achievements of the school children. One paragraph of the report (1) is: "The questions in grammar are the best proof that scholars may parse technically and point out the relations of words, their mood, case, person, number and gender; and yet, in the very sentences which they make use of to express these relations, and in quoting rules in justification of what they write, be continually making blunders; and may parse their sentences grammatically in the most ungrammatical language."

Here we have evidence that formal training in English grammar is not having much, if any, effect upon English composition, because pupils can "parse their sentences grammatically in the most ungrammatical language."

The Psychological Attack. — The decay of the old faculty psychology and the rise of modern psychology with its greater interest in practical problems led eventually to the psychological attack on the problem of transfer. The problem now becomes quantitative and the question is, "How much influence does improvement in one activity have on some other activity?"

We must, therefore, have an initial test of an activity, a period of practice in a second activity and a final test in the first activity to see if any improvement has taken place in this activity because of the practice in the second activity. The scheme for this experiment is roughly as follows:

Experimental Group	: Initial Test	: Practice	: Final Test
Control Group	: Initial Test	: No Practice	: Final Test

You will notice that there are two groups, a control as well as an experimental group. This control group has no practice period, but takes the initial and final tests, and is supposed in all other ways to be subjected to the same influences as the experimental group. Why is the control group necessary? Because if the experimental group made a large gain on the final test after the practice period, we would not know whether this large gain were due to the specific practice or to some other influence. The control group will check this, because if the gain is due to some other influence the control group will also show it. The real influence of the specific practice will therefore be the difference in gain made between the control and the experimental groups. This is the actual influence of the practice period, and is called the residual gain.

It will be obvious at once that our results will be more accurate the more alike the control and experimental groups are. Hence it is customary to equate these two groups, to make them as nearly alike as possible. In careful modern experimental work with school children the control and experimental groups are equated with reference to chronological age, mental age, grade, educational achievement and so on.

If, after all these precautions have been taken, the experimental group makes a greater gain in the final test than the control group, we have a quantitative expression of the amount of transfer of one activity upon another.

Results of Laboratory Studies. — The first experiments in transfer were those made by psychologists in their laboratories and were not directly connected with school subjects. Many of them were not as carefully made and controlled as we have suggested they should be in the preceding paragraph. Let us, however, see what general results were obtained.

Table 4 gives a very brief summary of some of these earlier experiments carried on in the laboratory. The results are grouped according to the kinds of functions

TABLE 4

LABORATORY EXPERIMENTS IN TRANSFER OF TRAINING

Author	*Date*	*Training*	*Transfer to*	*Result*
Memory:				
James	90	Memory passages	Other memory passages	Slight or none
Ebert and Meumann	05	Non-sense syllables	Numbers, letters, words, poetry, etc.	22% residual gain acc. to Dearborn

TABLE 4—*Continued*

Author	*Date*	*Training*	*Transfer to*	*Result*
Fracker	08	Memory of sound intensities	Similar stimuli Dissimilar stimuli	16% residual 3% residual
Sleight	11	Poetry, prose, tables	Dates, syllables, etc.	3–4% residual
Winch	08	Poetry	Historical prose	10% residual
Reaction-Perception:				
Gilbert and Fracker	97	Reaction to sound	Reaction-sound touch, color, etc.	52% No controls
Thorndike and Woodworth	01	Areas, weights, lines, letters, etc.	Similar stimuli	52% down to zero
Coover	16	Marking words	Other words	75% as much as in training
Kline	09	Cancellation of e. s.	Nouns, verbs	Interference. Less than control group
Bennett	07	Discrimination of shades of red	Other colors	60–23% gain
Coover and Angell	07	Sound intensities	Brightness intensities	10% gain
Sensori-Motor:				
Judd	08	Theory of refraction	Target under water	Improvement when depth is changed
Webb	17	Maze	Other mazes	19–77%
Coover and Angell	07	Sorting Cards	Typewriter reactions	Very little or doubtful

investigated. Under memory we have five experiments dealing with this type of reaction. The first line tells us that James in 1890 conducted some experiments. He asked the question whether training in one kind of memorizing would affect, or influence or transfer to other kinds of memorizing. He was raising the general question as to whether the memory can be trained. So he tested himself on a certain passage of poetry (initial test), then he practiced memorizing another kind of poetry (training or practice period), then he tested himself again on the first kind of poetry (final test), and found that no improvement had taken place. He repeated these experiments with some of his students and found slight gains, so in the "Result" column of Table 4 we find "slight (improvement) or none." These early experiments of James were carried out in what we would consider now-a-days a rather perfunctory manner. And so we must not lay too much emphasis on them.

Take the next experiment which was more carefully conducted by Ebert and Meumann in 1905. Here the training was in memorizing non-sense syllables and the question raised was whether practice in memorizing this material would increase one's ability in learning numbers or letters or words or poetry. Initial and final tests were given, but there was no control group, so Dearborn at a much later period worked over the tests again with a control group and estimated that the real amount of gain due to the training period (residual gain) was about 22 per cent.

In a similar way we might discuss the results of each one of the experiments listed in the table, but this is unnecessary. A general study of the table will be sufficient. First we note that the kinds of reactions in which the various subjects were trained are rather special and definite. The functions trained are narrow rather than broad. We find training in such narrow functions as estimating areas, lengths of lines, running mazes, sorting cards, while the broadest function is learning poetry. We do not find such broad functions as learning Latin or studying law or engineering. And again we note that activities tested for transfer effect are also specific and narrow, such as discriminating colors, typewriting reactions, memorizing dates or prose. We do not find the experimenters asking whether any specific kind of training influences the study of French or mathematics or study in general or memory in general or one's ability to be a doctor or lawyer. The reason why narrow functions have been chosen in the laboratory is of course clear. Such narrow functions can be more easily measured. Broad functions are difficult to measure accurately.

Another thing we note from the table is the great amount of similarity between the training function and the transfer function. There is similarity in the functions found in columns 2 and 3. Memory for non-sense syllables is paired with memory for numbers, etc.; marking certain kinds of words with marking other words; discriminating shades of red with discriminating other colors; learning some mazes with learning other mazes and so on. We do not find the question raised as to whether training in one function results in transfer to another widely different function, for example, whether training in running a maze influences typewriting reactions, or whether marking words influences reaction to sound or color; and so on. In other words we find in general the experiments are limited to narrow measurable functions and the transfer effects to similar narrow functions are measured.

When we run down the column headed "Results," we note that in general there is an appreciable amount of transfer effect. Almost all experiments find some transfer effect and several find a considerable amount. Hence we may sum up by saying that the fact of transfer is obvious; that the amount of transfer is considerable with closely allied functions; that the amount of transfer decreases as the functions become more and more unlike. If the functions are widely separated, much more so than any in our table, there is likely to be little or no transfer.

Results of School Experiments. — For educational purposes the laboratory studies we have just described are not particularly helpful, but they have established the general method of experimentation and the general fact of transfer under certain conditions. What shall we find with school children working with the broader functions of school subjects? Table 5 (p. 276), gives us a summary of the chief experiments. Most of these experiments were done on experimental and control groups of school children working under class room conditions. Under foreign languages the differentiation is generally between students studying Latin and those not studying Latin. Rarely have the control and experimental groups been equated for intelligence and this is important when comparing Latin and non-Latin groups, because of the selective influence of the Latin curriculum. Students studying Latin have been found to score slightly higher on intelligence tests than non-Latin students.

Keeping in mind the shortcomings of much of the experimental work, we may study the results. In three cases where the residual gain is high (29 to 34 per cent), we notice that the functions are very closely allied. Thus mental multiplication is closely allied to the other fundamental operations in arithmetic. In all other cases the transfer is slight or so small as not to be measurable. The study of formal grammar has very little or no influence upon correct usage of English or ability to reason about words and definitions. You remember the Boston Report in 1845 found that scholars "may parse their sentences grammatically in the most ungrammatical language." The study of Latin does seem to help a little in

TABLE 5

SCHOOL EXPERIMENTS IN TRANSFER OF TRAINING

Author	*Date*	*Training*	*Transfer to*	*Result*
Arithmetic:				
Starch	11	Mental multiplication	Adding, dividing, etc.	29% residual
Winch	10	Computation	Reasoning	Slight or doubtful
Grammar:				
Briggs	13	Formal	Reasoning	None
Starch	15	Formal	Correct usage	Slight
Foreign Languages:				
Swift	06	Latin and non-Latin groups	Spanish	Latin most progress
Starch	15	Latin	Size of English vocabulary	3–4% in favor of Latin students
Perkins	14	Latin	Spelling, definitions, etc.	Latin superior
Starch	17	Foreign language	Writing and reading tests	Transfer (no transfer for memory tests)
Science:				
Hewins	16	Biology	Biological tests	34% residual
			Non-biological tests	5% residual
Rugg	16	Geometry	Geometrical tests	32% residual
			Non-geometrical tests	7% residual

studying Spanish, in English vocabulary and spelling, but certainly not very much, and certainly not enough to justify the teaching of Latin for such transfer values alone.

If we sum up these school experiments, we find that they bear out the conclusions derived from the laboratory experiments. Closely allied functions show a fair amount of transfer, but as functions become less allied the amount of transfer dwindles and soon approaches zero. This takes us back to our discussion of learning. We learn what we practice. Learning is reacting. We learn to react to certain stimuli. Unless these stimuli appear in the new situation, we have not learned how to react to that new situation; there is no transfer. If the stimuli are somewhat like the old stimuli, we may be able to react to them; there is a little transfer. If they are very similar to the old stimuli, we know how to react; there is much transfer.

The effect of changing the stimulus to which we have learned to respond has been shown very effectively by Thorndike (4). In a class of college students only 6% made errors when called upon to find the square of $x+y$; but 28% made errors when called upon to find the square of b_1+b_2. If the stimuli are still further changed, Thorndike shows that the lack of transfer becomes still greater.

More Recent Studies of Transfer. — Several recent studies have dealt with the influence of the study of Latin. These show that after one year's study of Latin a gain of 1.8 score points on the Thorndike-McCall Reading Test is made by the Latins above the gain made by the non-Latins. On a vocabulary test Latin students in a year gained roughly five words of Latin origin, while non-Latin students gained only two; but both groups made the same gain in words of non-Latin origin. The study of Latin then was only helpful for English words of Latin derivation. In spelling English words of Latin origin Latin pupils gained one word more than non-Latin pupils. All of these and similar investigations merely confirm the earlier experiments. A study of Latin does transfer, it does help in building up an English vocabulary, in learning to spell English words, but the amount of such transfer is very small. Symonds (3) shows that if we want to build up a vocabulary, the thing to do is to teach words directly and the amount of gain so made is much greater than any to be expected indirectly through the study of Latin.

Latin should not be taught for these indirect values. They are too slight, too evanescent to justify the amount of time and labor spent on the study of the subject. Latin must justify itself on other grounds. It has other grounds. It should abandon the appeal to formal discipline. What we say about Latin is equally true of any other subject that rests its case on disciplinary values.

General Ability to Reason. — But we have not yet finished with disciplinary claims. Even if the gains for any subject are slight in such concrete things as English vocabulary, correct use of English and so forth, there may be gains in general ability to reason, in ability to generalize, to draw valid conclusions and other similar higher capacities. In fact such are the claims made in greater or less degree for Latin, for mathematics, for science and indeed for many other subjects. Which subjects as taught in the present-day high school show the most gain in general reasoning ability? Are there any subjects that show greater gains than others? In an extensive investigation covering 8564 high school pupils in Grades IX, X and XI, Thorndike (5) has answered these questions. As initial and final tests he used tests of selective and relational thinking, generalization and organization. The practice period was a year's work in high school subjects taught in the usual way. The gains in score in the final tests were then computed for each type of study. French, for example, gained 2.5 score points, i.e. taking French "increased a pupil's gain by 2.5 score points over what it would be if he had taken an average mixture of studies instead of French." French, as a matter of fact, led the list followed by chemistry and trigonometry. Latin comes a little further down along with general science and bookkeeping. Towards the bottom of the list come stenography, economics, cooking and sewing. But we shall not print the list because the differences between subjects are indeed very slight and not reliable, and a second similar investigation might change the places of many of them. For, as Thorndike says, "the results are in pronounced opposition to the traditional view that certain subjects produce much more general improvement in ability to think than others." Latin then is not any better in this respect than French, nor is algebra better than bookkeeping.

Another very important conclusion emerges from Thorndike's investigation, namely, the value of general intelligence in making further gains in ability to think and reason. After one year's work in high school the highest one per cent in general ability gained 20.5 points, while the lowest one per cent gained 1.5 points, regardless of the studies they had taken. It is not the studies, then, that cause differences in gain, it is the abilities of the students taking these studies. "Those who have most to begin with gain the most during the year. Whatever studies they take, they seem to produce large gains in intellect." And Thorndike's final conclusion, "The intellectual values of studies should be determined largely by the special information, habits, interests, attitudes and ideals which they demonstrably produce. The expectation of any large differences in general improvement of the mind from one study rather than another seems doomed to disappointment."

Transfer of Moral Qualities. — Most of our evidence for transfer rests upon experimental work with school studies. There are, however, a few studies dealing with the transfer of moral or non-intellectual values. These all point in the same direction as the studies dealing with intellectual values. Honesty learned in one situation cannot be expected wholly to transfer to a different situation; the greater the difference the less will be the transfer. The teaching of good citizenship in school must not be expected to transfer wholly to out-of-school situations. We may learn to be unselfish and play for the team in baseball or football, but it may affect very slightly our attitude when away from the "play-ground" situation.

Method of Teaching Affects Transfer. — We must not forget that the amount of transfer we may find in any given school situation will be determined by the method of teaching. We cannot say that Latin will transfer 10 per cent to English vocabulary; or if we say this, we mean Latin as generally taught, or Latin as taught in a specific way. By changing our method of teaching we may change the amounts of transfer effect to other subjects. Any teacher could easily increase the amount of transfer effect on English vocabulary by teaching Latin with special emphasis on and reference to growth of words. All this is very obvious. In our psychological terminology, what we are doing in altering the methods of teaching a subject is making it more or less allied to other subjects. Emphasizing Latin words and their evolution in other languages is making this particular kind of Latin more allied to English vocabulary, and hence as we have seen from the psychological experiments in the laboratory we are to expect more transfer effect. We learn exactly what we practice. If we practice etymology, we learn etymology, and where Latin and English etymology overlap the one helps the other or in other words we have transfer. The method by which a subject is taught will determine its transfer value. Beware, however, of teaching Latin in order solely to teach English Better far teach English directly.

The Mechanism of Transfer. — How does transfer take place? We have suggested this in all that has gone before. We spoke about subjects overlapping, being closely allied, having things in common. The more alike two stimuli are, the more likely are they to call forth the same reaction. "Ratio" is more like "ratiocination" than "vernunft" is, so that a knowledge of "ratio" is more likely to help us in understanding "ratiocination" than a knowledge of "vernunft" is. Bicycling is more like driving an automobile than skating is, hence knowledge of how to ride a bicycle may help us more in learning to drive an automobile than knowledge of how to skate. The stimulus of mathematics work before one with the reaction "must do it well," "must be accurate," is more like the stimulus of chemistry work before one than it is like the stimulus of English composition, and hence we might expect more transfer of reaction "must do it well," "must be accurate" to one's chemistry work than to English composition. But we would not expect much transfer in any case, unless the total situations of "doing arithmetic" and "doing chemistry" were much alike.

The common elements in two situations determine the amount of transfer from the one situation to the other. These elements which are common to two situations may be identities of content or identities of procedure, and the method of teaching will determine which identities are most emphasized. Latin may be taught with an emphasis upon the words which are like English words (identities of content), or it may be taught with an emphasis upon accuracy in thought and expression in general (identities of procedure).

Of course we do not react to the common elements alone. We react to the total situation always. And methods of training may make us see or appreciate identities in certain situations which other methods do not, and so some methods obtain more transfer than others, because they emphasize identities in certain directions. But most important of all evidently for general transfer is the general ability of the individual who is reacting. If one possesses much general ability he is likely to note and appreciate more identities than if he has little general ability. If intelligence is the ability to react adequately to a novel situation, then we may say that the intelligent individual is able to do this better, because he is more likely than the unintelligent to see the identities between the new situation and previous situations and hence make an adequate transfer to the new situation. He solves the problem because it is like a problem he has solved before. He sees the identities of content or procedure, which the less intelligent individual fails to see.

Application to Educational Practice. — No subject should be kept in the curriculum for its transfer values alone. It must have other and independent claims and its place in the curriculum must be considered solely in the light of these. If some subjects resulted in much greater general improvement in ability to think than others, they surely should receive greater consideration. But all the evidence is against this, so that any claim in this respect must be waived aside until established. Each subject, therefore, must stand upon its own feet and be taught for its own sake. Do not teach a subject in such a way as to attain large transfer effects to another subject. Teach the other subject directly and you will attain better results with less time and less labor.

Be very sceptical of the general moral value of religious education or of the general character-building qualities of physical education, until these claims have been more definitely demonstrated. If the battle of Waterloo was won on the playing fields of Eton, where was the battle of Austerlitz won? If the college athletes play fair on the football field is that any guarantee that they will play fair in the class examination?

In school administration do not jump to the conclusion that a good first grade teacher will naturally make a good sixth grade teacher or a good elementary teacher will make a good high school teacher, or a good teacher will make a good principal and so on. The recognition of general goodness is not enough.

Summary

1. The problem of transfer is the problem of reciprocal modification. What influence does a given modification have upon other modifications?
2. The world at large has learned to assume transfer in some cases, but not in others.
3. Formal discipline is the belief in a very general and widespread transfer. Belief in this theory of formal discipline has influenced educational practice very considerably.
4. If the utilitarian value of a subject is attacked, it generally lays claims to disciplinary values.
5. Belief in wide transfer is still common among many educators and much that is taught is still defended on the grounds of transfer values.
6. Psychological investigation of transfer values has introduced a definite experimental procedure, with initial and final tests flanking the practice period, with control and experimental groups equated.
7. The results of laboratory studies show considerable transfer between narrow allied functions; but little transfer between other functions.
8. The results of school experiments dealing with much broader functions show small transfer effects. There may be transfer from one school subject to another. The fact of transfer is obvious. The amount in general is very small.
9. No one high school subject, as at present taught, is markedly better than any other in increasing a pupil's general ability to reason, to generalize, to carry on abstract and relational thinking.
10. The most important factor in the increase of general ability to think is the general intelligence already possessed by the student and not the particular course of study he may take.

11. The transfer of moral qualities seems to follow the same trend as the transfer of intellectual qualities. In general we greatly over-estimate the amount of transfer of such qualities.
12. The method of teaching affects the amount of transfer to any given situation. It may do this by emphasizing common elements of content or common elements of procedure.
13. The mechanism of transfer is by means of the elements common to two situations. And these elements may be either of content or of procedure and the amount of similarity perceived is determined by the method of teaching and the intelligence of the learner.
14. In school we should never rely on the transfer values of a subject. To pin our hopes on much transfer is to court disappointment.

Review

True-False Statements

As a review mark the following statements true or false:

1. It is doubtful whether any particular subject rather than another is an effective instrument for improving generally the mind's ability to think.
2. Individual differences in transfer of training depend upon degree of intelligence.
3. The study of English grammar leads to improvement in written composition and in speech.
4. The residual gain in a transfer experiment measures the real amount of transfer.
5. Teaching the theory has never been found to cause improvement in the actual practice of a function.
6. If you wish to improve your memory, it is well to practice memorizing lists of numbers or dates.
7. The study of mathematics increases one's general accuracy.
8. Training in addition may show transfer to multiplication.
9. Training in Latin helps one in his knowledge of English words.
10. There may be transfer of attitudes as well as transfer of content.
11. The method of teaching a subject will not influence the amount of transfer to other subjects.
12. The disciplinary value of Latin justifies its retention in the curriculum.
13. Any change in the concrete particulars reasoned about will interfere somewhat with our reasoning.
14. Pupils having different amounts of mental ability tend to make about the same amount of improvement on general reasoning tests after one year of high school work, regardless of the course of study they follow.
15. Spread of improvement diminishes as functions become more unlike.
16. The increase in knowledge of English words through the study of Latin is an example of transfer by means of identities of procedure.

17. To promote a professor, distinguished for research, to the position of dean indicates a belief in a great amount of transfer.
18. Underline the best response:
    Transfer of training from one function to another would seem to take place by means of: (1) generalization; (2) mental exercise; (3) identical elements; (4) ideals.
19. Check the best response:
    The most important factor in general mental improvement regardless of the subject studied is:
    (a) mental set.
    (b) intelligence.
    (c) associative shifting.
    (d) concentration.

Advised Readings

Starch, D.: *Educational Psychology*, New York, 1927, Chapter 14.

Thorndike, E. L.: *Educational Psychology: Briefer Course*, Teachers College, Columbia University, New York, 1914, Chapter 18.

References

1. Caldwell, O. W., and Courtis, S. A.: *Then and Now in Education*, Yonkers, 1924, p. 52.
2. Mais, S. P.: *A School Master's Diary*, London, 1918.
3. Symonds, P. M., and Penney, E. M.: "The Increasing of English Vocabulary in the English Class," *Journal of Educational Research*, 1927, 15, 93–103.
4. Thorndike, E. L.: "The Effect of Changed Data upon Reasoning," *Journal of Experimental Psychology*, 1922, 5, 33–38.
5. Thorndike, E. L.: "Mental Discipline in High School Studies," *Journal of Educational Psychology*, 1924, 15, 1–22 and 83–98.

CHAPTER XIII

CONTINUOUS EXERCISE WITHOUT REST

Fatigue. — So far we have always assumed that in trying to make any improvement, the exercise has been carried on for given periods with rest, change of occupation and sleep intervening. We now raise the question as to what happens to a learning or improvement curve, if the exercise is continuous. In other words how long can a person work at a task without showing decrease in efficiency? If we practice anything continuously without stopping, will the improvement curve continue to mount indefinitely? Of course we know it will not. Something happens that we call fatigue. At any rate individuals stop working relatively soon, because they become restless, bored, uncomfortable and they seek other occupations or rest. The length of time any given individual will work continuously at any task varies enormously according to his interest in the task, his desire to attain a given end, the external pressures and compulsions holding him down to the task and the like.

Mental and Physical Fatigue. — Mental and physical fatigue are two names for the same phenomenon. There is no real difference between them. Fatigue is something that lowers the efficiency of the organism for any kind of work. If the work has mainly to do with the large muscles of the body, as in walking or sawing wood or playing golf, the discomfort arising from continuous exercise is called physical fatigue. If, on the other hand, the work involves none of the larger muscles and is mainly concerned with thinking or reacting with speech or subvocal movements, as in multiplication or studying Latin or writing a dissertation, the discomfort arising is called mental fatigue. There is no sharp line dividing the one from the other. The educational psychologist is mainly concerned with the latter type of reaction, and decreasing efficiency due to continuous exercise of this sort is what we shall call mental fatigue.

Measurement of the work of a worker is difficult. There are two ways in which this can be measured. The first and simplest and most objective is to measure the quantity and quality of the product produced, as in tasks in arithmetic we may measure the number of tasks attempted and the correctness of these tasks. The second way is to consider the amount of effort put forth by the worker, the discomfort and obstacles overcome by him. One pupil may work harder than another, but produce less work. This is obvious, and in school the teacher may take cognizance of this so far as he can. He may give a dull child credit for effort put forth. Indeed, the A.Q. technique is an attempt to provide a measure of this kind, for it tries to evaluate the product in terms of the general intelligence of the child. Two children producing the same amount and quality of work may have very different A.Q.'s. This same work may be good with reference to one child's capacity, but poor with reference to another's.

The Laboratory Experiments. — Let us glance now at the results of the measurement of fatigue in the psychological laboratory. Table 6 gives a condensed summary of many experiments arranged in chronological order. Notice that the function tested or the work done is restricted to the type of reaction which we popularly say involves a lot of thinking. The reactions called for are such as involve the finer musculature of the body. They do not involve much overt muscular movement. They are more mental than physical, because we are interested in the effects of mental fatigue. The arithmetical operations of adding and multiplying are the most commonly used reactions. Notice also that the duration of time is generally rather short. This is particularly true of the earlier experiments. The later experiments deal with longer time periods and we shall see presently the reasons for this. The subjects used in the experiments are generally few in number and in several cases the subject is the experimenter himself. It is difficult to obtain subjects who will consent to undergo long, monotonous and hard work.

Now if we read down the column of results, we are first of all struck by the contradictory character of the results. Some experimenters report a considerable fatigue effect and others none. One hour of addition is reported as showing considerable fatigue whereas another experimenter reports no fatigue effect. Some of this discrepancy is due to the elusive character of the thing that is being measured, the differences in the amount of interest in the task, and the differences in the amount of effort put forth. If a subject knows that the experiment is for the purpose of obtaining a measure of fatigue, if the subject believes that mental fatigue shows itself quickly, he is likely to be unconsciously influenced in his attitude towards the work. We can readily produce a fatigue curve if we are anxious to do so.

The earlier experiments were more or less ambiguous as to the rapid onset of fatigue. Thorndike's first experiment in 1900 with work periods of 3 to 8 hours showed no measurable fatigue results. This made him sceptical of the fatigue produced after short periods of work. Hence when he returned to these fatigue experiments in 1910 and 1911 we note that he lengthens the work period and also makes the work more difficult, i.e. by introducing mental multiplication of three place by three place numbers. When he does this he begins to find slight but definite fatigue effects. This work of Thorndike cul-

TABLE 6

SUMMARY OF LABORATORY FATIGUE EXPERIMENTS

Experimenter	*Date*	*Reaction*	*Duration*	*Subjects*	*Result*
Burgerstein	91	Adding multiplying	1 hour	Children	None; perhaps gain
Bergstrom	94	Translating German	4 hours	Self	None
Holmes	95	Addition	36 min.	Children	Doubtful or none
Bettman	95	Reactions	1000 trials	Adults	Slight increase in time
Oehrn	95	Counting, memory, adding, writing, reading	2 hours	10 adults	Very slight fatigue effects. Interference in nonsense syllables
Weygandt	97	Addition	90 min.	3 adults	Considerable. One fourth better after rest
Vogt	99	Addition	90 min. daily	Self	Considerable. 24% better after rest
Vogt	99	Memory for digits	75 min. daily	Self	9% better after rest
Thorndike	00	Cancellation, areas, memory, complex mental work	3 to 8 hours	Woodworth	None

TABLE 6—*Continued*

Experimenter	*Date*	*Reaction*	*Duration*	*Subjects*	*Result*
Lindley	00	Addition	1 hour	3 adults	None
Bolton	02	Addition	1 hour	Adult	Considerable decrease
Kafemann	02	Addition	90 min. 8 days	Self	2% better after rest
Henmon	04	Addition	1 hour	Adults	None
Specht	04	Addition	—	Adults	Marked fatigue even after 10 min.
Whipple	10	Cancellation	2 hours	Self	None
Thorndike	11	Multiplication 3 place × 3 place	4 to 12 hours	Adults	21–59% more time than after rest
Thorndike	12	Addition	1½ to 2 hours	—	Fatigue effect of 6%
Arai	12	Multiplication 2 place × 2 place	2 hours	Adults	Increased efficiency
Arai	12	Multiplication 4 place × 4 place	12 hours	Self	Increase in time—119%. efficiency equals 75% of initial
Starch & Ash	17	Addition	2 hours	Adults	4.3% loss
Ritchie	24	Multiplication	10 hours	Adults	Slight drop in number done. Marked rise in errors

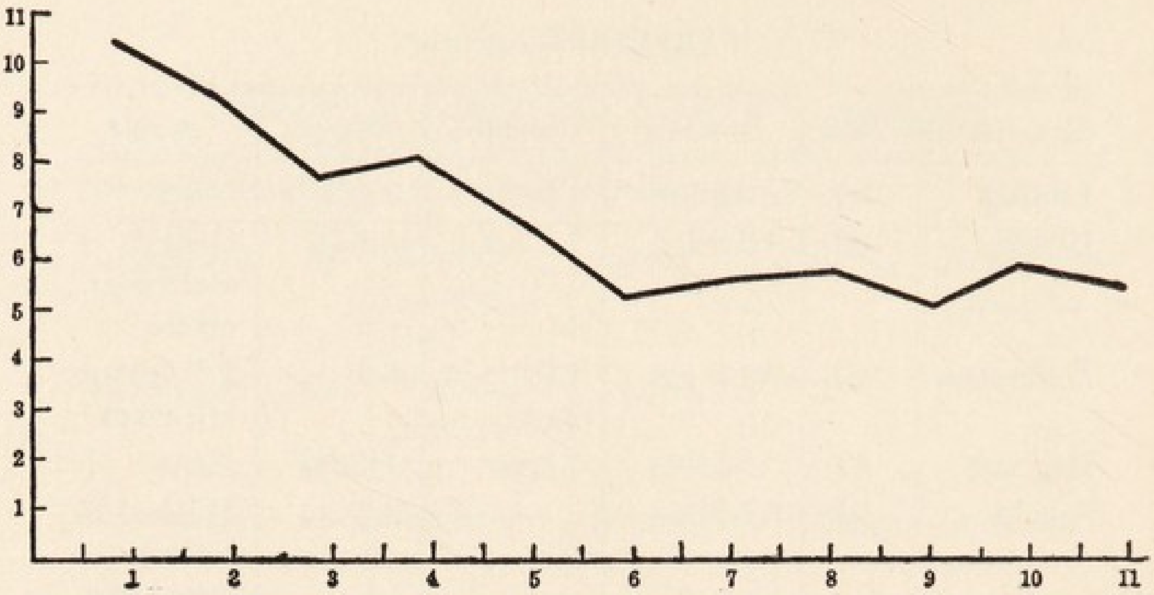


FIGURE 69. — Work Curve for Mental Multiplication. The vertical axis represents the number of 4-place by 4-place examples worked per hour of working time. The horizontal axis represents the successive hours of work from 11 A.M. until 11 P.M. (Drawn from date of Arai, *Columbia Contributions to Education*, No. 54, after Robinson.)

minates in the work of his student, Arai, who still further increased the difficulty of the work and the length of the work period. Arai worked continuously for twelve hours on four successive days at a hard mental task, namely the mental multiplication of four place by four place numbers. She reduced as far as possible every physical reaction. She merely looked at the card with the two series of numbers until she had memorized them, so as to avoid eye strain. Then she made the mental calculation and when this was finished she wrote down the answer. In this way such physical reactions as writing and looking were reduced to a minimum. After twelve hours of this severe mental application, the subject was still able to multiply although she was working much more slowly and making a great many more errors. Figure 69 shows the number of examples worked per hour during the twelve-hour period. The number done decreases somewhat at first, but towards the end remains surprisingly constant for many hours. At the end of this long period she was still able to multiply. She was not fatigued to the extent of not being able to multiply anything at all. Thorndike estimates her probable efficiency to be about 75% of her total multiplying ability. Evidently, then, mental fatigue does not arise as readily and as quickly as is popularly supposed. Given an adequate motive, an individual can work for many hours at a mental task without loss in efficiency and for very many hours without great loss in efficiency. After twelve or sixteen hours the problem will become complicated by the onset of sleepiness. But within the ordinary waking period we may justly say that it is hard to obtain real mental fatigue under experimental conditions.

We will note, however, the short-comings of our experiments, particularly as regards the type of work. This always is a more or less mechanical type of work, doing the same type of thing over and over again. This is chosen so that efficiency from hour to hour may be accurately measured. Would fatigue show itself earlier and more definitely with more creative work, such as planning a campaign, writing a book, composing music and the like? It is impossible to answer this. Such things are hard to measure as they are being done from hour to hour. Furthermore, twelve hours spent in writing poetry is a mixture of many different kinds of activities, and how much time is involved in pure creation is doubtful. Again from accounts of authors, scientists, composers, who have worked continuously for long periods of time when the *divine afflatus* was upon them, we may surmise that the amount of mental fatigue would not likely be greater than such as would arise from the very dry and monotonous task of multiplying mentally four place by four place numbers. The reader might well try how long it takes him to multiply 6483 by 2795 without using pencil or paper. If pure creative work could be isolated under laboratory conditions and if a worker could be made to work at it continuously, without dropping down from the creative level, it is possible that fatigue would show itself earlier than it does in hard but routine mental work.

To sum up the results of the laboratory experiments, we may rightly say that mental fatigue is slow in onset and that an individual can work several hours without rest without any appreciable loss of efficiency. The usual restlessness which arises after a few hours of mental work is not mental fatigue, but discomfort, boredom, desire to do something more attractive.

The School Experiments. — Fatigue in adults in the laboratory may be hard to register. The situation may be different with children in school. Does not the ordinary work in school tire the child? Does not mental fatigue increase during the school day and leave the child less efficient at the end? Several experimenters have raised this problem and the results of their work are summarized in Table 7. The usual procedure is to give similar tests at the beginning of the school day and at the end, and then to measure any difference that may be present. Sometimes similar tests are given at different periods during the school day for several days and efficiency at 9 o'clock, 10 o'clock and so on up to the last period measured.

Now if we read down the column headed "Results" in Table 7, we are struck by the fact that little or no decrease in efficiency is reported. Evidently children are able to work (we do not say they actually do work) as efficiently at the end of the school period as at the beginning. The work of the ordinary school day does not decrease a child's efficiency in such "mental" work as multiplication, addition, dictation and the like.

Now this seems to contradict the experience of many teachers. They report that children are tired out at the end of the school day; that they cannot do "hard" sub-

TABLE 7

RESULTS OF SCHOOL FATIGUE EXPERIMENTS

Author	*Date*	*Tests*	*Results*
Sikorski	79	Writing from dictation	Errors increase slightly
Bolton	92	Memory for digits	No loss in efficiency
Laser	94	Addition and multiplication	Work improved (practice)
Friedrich	97	Addition, multiplication, dictation	Very slight decrease
Ebbinghaus	97	Memory for digits. Addition, multiplication. Completion exercises	No difference. No real difference. Decrease in efficiency (?)
Ritter	00	Marking letters. Dictation	No difference.
		Dictation of unrelated words	Slight decrease
Thorndike	00	Adding, multiplying, cancelling, memorizing, etc.	No difference between late and early work
King	07	Addition, multiplication, judgment	No difference
Heck	13	Arithmetic	Increase in amount; decrease in accuracy
Dawson	23	Multiplication	No difference

jects, such as arithmetic as well as they can in the forenoon. The psychologist believes that what is happening here is not fatigue at all, but physical restlessness due to constraint, boredom due to uninteresting subject-matter, increasing desire to play as the time for play draws near. Many a teacher has been surprised to notice how efficiently children can and will work even at the end of the day when the thing they are doing really interests them. There is no question of fatigue then. Some of the more progressive schools report a difficulty of getting children to go home, so much do they love their school work. No question of fatigue there. By monotonous senseless drill, by uninteresting subject matter, by stupid teaching we may bore our pupils and drive them to hate school and all learning, but even then we are not likely, during the few hours of school each day, to occasion mental fatigue. The chances are that few, if any, students in school or college ever become mentally fatigued.

Nervous Breakdown. — It is frequently said that this student or that student has had a nervous breakdown because of studying too much, because he is mentally fagged out. In all such cases it can readily be shown that mental fatigue was not the cause of the breakdown. In fact in most cases the amount of study done is woefully far from producing mental fatigue. It is not the study that produces the breakdown, but the worrying about the study. There are as many mental breakdowns among indifferent students as among assiduous students. A mental collapse may be brought on by physical disturbances, by emotional up-sets, by fears and worries, but hardly by pure mental fatigue. Not study, but worry about study and more likely worry about something else causes the breakdown.

Feelings of Dissatisfaction. — Another interesting result of the psychological experiments on fatigue is the suggestion that the satisfyingness of a task is not directly correlated with the efficiency of the worker. We give up unsatisfying tasks if we can, not because they produce mental fatigue, but because they are unsatisfying. If we persist in spite of the unsatisfyingness, we are frequently as efficient as when the task was satisfying. Figure 70 shows an attempt to measure efficiency at a task and feelings of satisfaction at the same time. Several subjects worked for about five and a half hours at a mental task divided into 15 units. The solid line of Figure 70 shows that their work did not decrease in efficiency as a

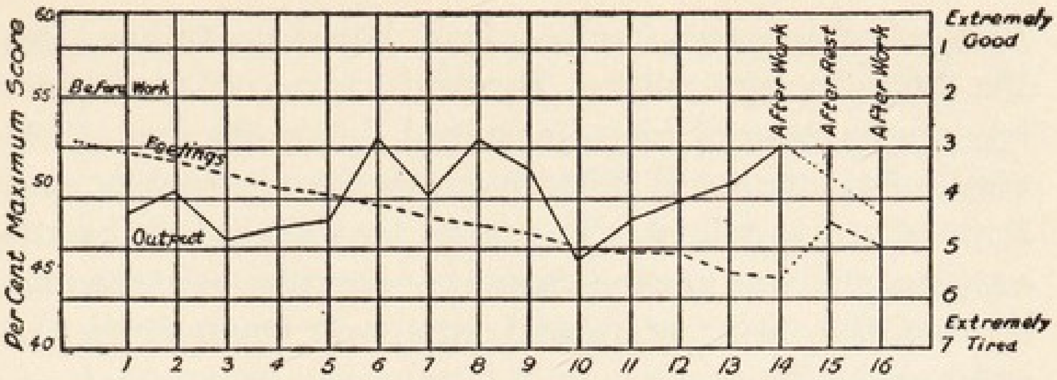


FIGURE 70. — Relation between feelings and output. (From Poffenberger's *Applied Psychology*.)

whole. After each unit they rated their feelings from "extremely good" to "extremely tired," and the broken line curve in the figure shows how more and more unsatisfying and boring the work became. They felt tired, but their efficiency was not decreased. If made to work, an individual may be just as efficient as when he works because he wants to. It is not a desirable state of affairs. But it should be a suggestion to us, so that we do not give up our own work as soon as we lose zest for it, and excuse ourselves with the rationalization that it is not efficient because it is unsatisfying. Although school work should as far as possible be made satisfying to the pupils, not all work in life can be so made. A pupil should learn to work at unsatisfying tasks, if necessary, and be shown that such work may be really efficient.

Suggestions for the School. — For school purposes then we should adhere to our general policy of placing the less interesting subjects at the beginning of the school day, not because pupils are mentally fatigued at the end of the day, but because they are less bored and less restless at the beginning. Notice that we lay stress on interest in a subject rather than hardness. A "hard" subject, like arithmetic, may be successfully taught and learned at the end of the day, if the pupils are vitally interested in it, and occasionally a teacher can be found who knows how to make arithmetic an interesting and vital subject.

Amount of interest, freedom from boredom, rather than the difficulty of a subject should be our guide in determining the order of subjects studied during the day. This order of interest will differ from teacher to teacher and from school to school. The point that the psychological experiments have made clear is that we need not take into account the factor of mental fatigue in determining the order of studies.

Intense Motivation and Fatigue. — Mental fatigue, producing a loss of efficiency in mental work, does not readily appear even after several hours continuous work under normal pleasant conditions. Intense motivation may postpone a drop in efficiency in mental work even under abnormal conditions. Knight and Remmers (3) have shown how the intense desire to make a college fraternity may keep men at a high degree of mental efficiency in spite of loss of sleep and unpleasant physical conditions. During the five day's initiation to a fraternity, several college students underwent severe physical exertion with only one or two hours sleep each night. At the end of this period they were given addition tests and told that unless they did well on these tests they would not make the fraternity. The motive for doing well on these tests was very strong. The tests consisted of 120 minutes continuous work with a record of each of the twenty-four periods of five minutes each. The results show a very slight loss in efficiency at the end of this two hour period. The average per cent correct for the first 12 periods was 85; for the last 12 periods 83. The total amount accomplished during the first 12 periods was 256 examples; during the last 12 periods 249 examples. Furthermore, these tired students under the pressure of this strong motive did twice as well as other college students taking the test under ordinary class-room conditions with no strong incentive. Evidently, therefore, mental fatigue is not only slow in appearing under ordinary conditions but it can be counter-acted by a strong incentive under very abnormal conditions.

Fatigue in Industry. — We cannot in this book go into the extensive investigations of fatigue which have been carried out in industry. We may note, here, however, that industry in general is not interested in the question of mental fatigue, but rather in the general problem of lack of efficiency, whether this be due to mental or physical factors or feelings of boredom, restlessness or any mixture of these. Physical factors are of much greater import here than in school work.

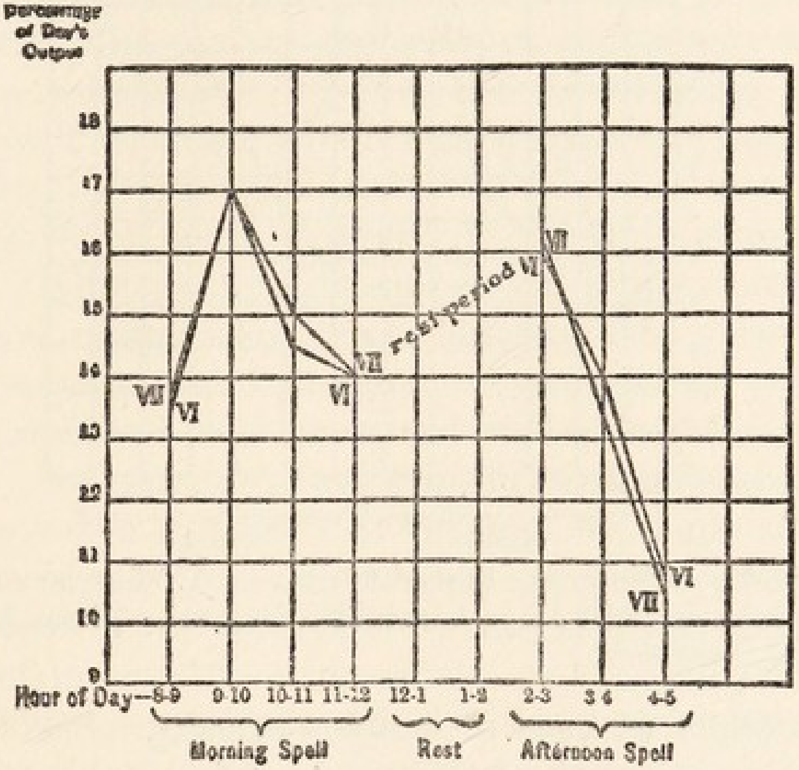


FIGURE 71. — Distribution of output over working day. Work of Italian typesetters. (From Muscio's *Industrial Psychology.*)

Measurements such as those shown in Figure 71 are of value to industry. The work curve rises in the forenoon and then gradually descends. During the afternoon the descent is continuous. Does fatigue, boredom, or loss of interest affect the attention and care a worker gives to his work? Figure 72 shows the number of accidents according to the time of occurrence during the day. The greatest number of these occur just before the noon-hour rest and just before the close of work at the end of the day. Fatigue, whether mental or physical, lack of interest and feelings of unsatisfyingness increase as the day wears on and the total effect of these factors leads to a greater number of accidents.

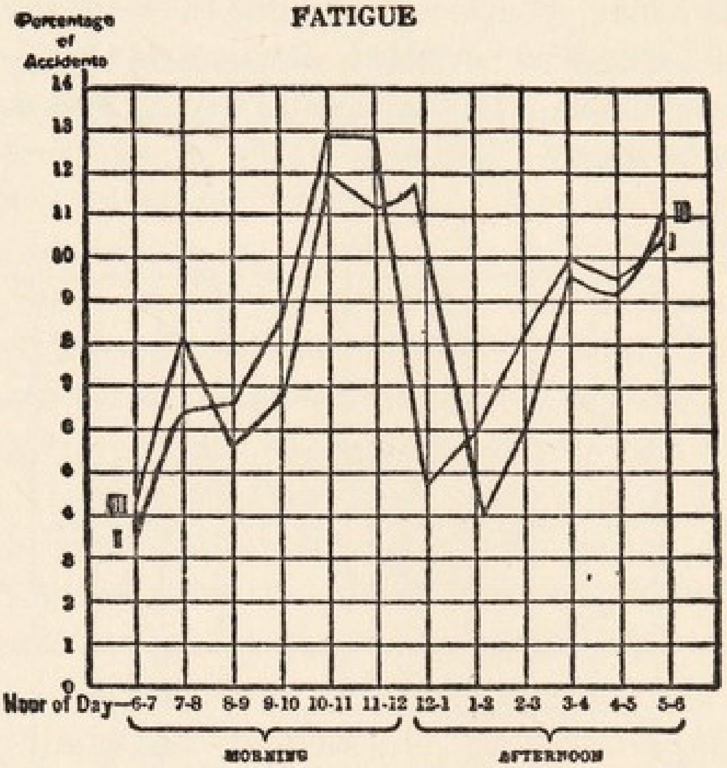


Figure 72. — Shows the time of occurrence of industrial accidents. I in Germany in 1887; III in Lancashire, England. (From Muscio's *Industrial Psychology.*)

The Effect of Loss of Sleep. — Closely allied to the general question of what happens when we work continuously without rest is the question of what would happen if this lack of rest is taken to include lack of sleep. Sleep is one of the ordinary phenomena of life that we take for granted, but just what it is, is very difficult to determine. There are numerous investigations, physiological and psychological, with regard to it, and there are many hypotheses as to how it is caused and as to what it really is. Into all these we cannot enter here.

We shall merely touch upon the question as to what happens psychologically when people go without sleep for longer or shorter periods.

With a group of college students it was found (5) that one night's loss of sleep was not sufficient to interfere with their efficiency on the ordinary intelligence test. Although many said they were tired, this feeling of tiredness did not make them do poorer work on the test. Even two nights' loss of sleep has been shown (4) not to cause any decrease of efficiency in mental work, although all the subjects report the necessity for putting forth greater effort after such a deprivation of their usual sleep. How much more loss of sleep an individual can sustain without showing a decrease of efficiency we do not know. Much will depend upon the subject's general state of health and what he is doing in the intervals between tests. The strength of the motive to work will also be a determining factor. What the experiments so far do demonstrate, is that one or two nights' loss of sleep is not a very serious disturbance for the individual. Evidently the organism has a great deal of ability to adapt itself to unusual conditions. In no sense, however, should these experiments be interpreted to mean that students can study all night without harm. This may be possible, but in general the safe rule is to get as much sleep as seems necessary.

Ventilation. — We all know that it is much pleasanter to work in a well-ventilated room, where the temperature is not too high and the humidity is not too great. If the room becomes hot and "stuffy," we begin to feel uncomfortable. We feel drowsy, disinclined to work and frequently we complain of headaches and other minor discomforts. In a badly ventilated classroom we notice the flushed faces of the children. They are frequently restless and not inclined to work. Experts tell us that optimum conditions in ventilation call for a temperature of about 68 degrees Fahrenheit, a humidity of 60 per cent and 45 cubic feet of outside air per person per minute. Nevertheless, it has been shown by Thorndike that individuals can do efficient mental work under much less desirable conditions, namely with a temperature of 86 Fahrenheit, a humidity of 80 per cent and with no fresh air or movement of air. Under such uncomfortable conditions good work can be done, and furthermore this can be kept up for several days. Manual work would seem to decrease in efficiency under bad ventilation quicker than mental work. We do not know what the effect on the organism would be if work were continued under such unfavorable conditions for several months or years at a stretch.

This ability of the organism to work effectively under adverse atmospheric conditions falls in line with its ability to withstand the effects of fatigue and loss of sleep. It is another indication of the great adaptability of the human body. We must not, however, draw the conclusion that ventilation is of no importance in school. It unquestionably is. But, on the other hand, we must not allow changes in temperature or uncomfortable atmospheric conditions to be brought forward as excuses for poor work. If such atmospheric conditions are not too unpleasant, we should neglect them and go on working with the belief that our work will not suffer because of them. Thousands of summer schools testify that students can do good work when the thermometer is high and the humidity great, although the conditions for work may not be as comfortable as during the regular semester.

These psychological experiments in ventilation must not be used as an excuse for badly-ventilated school buildings. It is more comfortable to work in a well-ventilated room with a moderate temperature and with abundance of fresh air, and we must insist upon optimum conditions as far as we can. It would be as unreasonable to conclude that ventilation does not matter, as it would be to conclude from our sleep experiments, that sleep does not matter and that an individual ought to make a habit of staying up all night every two or three days.

The Effects of Drugs. — Closely connected with the effects of fatigue and loss of sleep are the psychological effects caused by drugs.

Alcohol has a depressing effect upon the organism, contrary to the popular belief that it is a stimulant. According to the dose administered it retards in general the reactions of the individual. It causes a loss in efficiency in such tests as tapping, color naming and finding the opposites of words. Individuals vary greatly in their ability to tolerate alcohol. Hollingworth (1) found that resistance to the drug is positively correlated with general competence. The person of high general ability may possess a more stable nervous system, which is less readily affected than the less stable nervous system of the less competent individual.

Caffein in small doses is slightly stimulating and has been found to increase somewhat the efficiency of mental work. Large doses, however, soon lead to a loss of efficiency. Coffee may be the cause of sleeplessness. Ordinarily, however, this is caused more by expectation than by the coffee. We are unable to go to sleep because we expect not to go to sleep readily after drinking a cup or two of coffee. Hollingworth showed that loss of sleep was caused as much by giving syrup with no caffein as with doses of 2 or 4 grains of caffein, which correspond roughly to one or two cups of coffee. Larger doses of caffein consisting of 6 grains, or three cups of coffee, do interfere greatly with an individual's normal sleep habits.

Table 8 shows Hollingworth's comparison of the effect of alcohol and coffee. The plus signs indicate increase in ability on the tests and the minus signs indicate decrease. The figures indicate per cent of the work gained or lost,

TABLE 8

Comparing the Effect of Alcohol and Coffee on Various Tests (after Hollingworth)

	Bottles			Grains	
	3–4 Beer	*5–6 Beer*	*6–9 Beer*	*3–4 Caffein*	*6 Caffein*
Pulse	+ 8	+ 10	+ 19	——	——
Steadiness	—68	—241	—370	—2.0	—583.0
Tapping	— 7	— 13	— 14	+3.1	+ 3.8
Coordination	— 6	— 10	— 20	+0.3	— 4.0
Color Naming	— 2	— 7	— 12	+7.0	+ 4.4
Opposites	— 5	— 12	— 23	+5.1	+ 6.2
Adding	—10	— 15	— 16	+1.8	+ 2.5
Substitution	— 4	— 9	— 6	——	——
Memory	—21	——	— 60	——	——

based upon records taken without drugs. Thus we note that in all the tests, motor or mental, alcohol leads to decreases in efficiency. The only increase caused by alcohol is in the pulse rate. Coffee, on the other hand, leads to a slight increase in efficiency in most of the tests, the most noticeable decrease being in the steadiness test.

Tobacco has been difficult to experiment with, because it is difficult to introduce the drug without the knowledge of the subject. In all these drug experiments, it is absolutely necessary to disguise as far as possible the purpose of the work, because subjects are so much influenced by preconceived notions as to what ought to happen to them. One experimenter (2) has, however, overcome the difficulty with tobacco by means of an experimental pipe with electrically heated hot air. When blindfolded none of the subjects were able to detect the difference between the experimental pipe with hot air and the real pipe with tobacco, while someone else was smoking in the same room. The immediate effect after smoking a pipe of tobacco on smokers seems to be slightly favorable in such mental operations as addition, where old habitual bonds are functioning; but the effect seems to be unfavorable for learning, where it is a question of the formation of new bonds. Non-smokers are affected unfavorably in most tests. It would seem, therefore, that such mild drugs as coffee and tobacco, as usually taken, exert very little influence on mental operations.

We have no authentic experiments as to the influence of tobacco or coffee on children. Presumably, such effects would be less favorable than with adults. In general, therefore, such drugs should not be given to children. As children grow older the effects of such drugs, as far as we know them, should be told to the adolescent. He should not be frightened by monstrous stories of the evil effects of tobacco, for example. Such exaggeration very readily defeats its own purposes. He will soon find out the untruth and his confidence in the teacher as an authority in other matters will be undermined. Tobacco may not be very good for high school boys, but it may not be very bad. It is, however, undoubtedly much better for them to smoke openly in public than to disguise their practice by lying and subterfuge.

Summary

1. There is no real difference between mental and physical fatigue.
2. All work, if carried on continuously without rest, tends to become less and less efficient, but the point at which loss of efficiency begins is often difficult to determine.
3. Feelings of dissatisfaction or of boredom and the like are not good measures of the onset of fatigue or of decrease in efficiency.
4. Laboratory experiments in fatigue show us that an individual may work several hours at a difficult task without serious loss of efficiency.

5. The school experiments in fatigue show us that loss of efficiency in mental tasks does not develop because of fatigue during the ordinary school day.
6. It is not necessary to consider the possible fatigue effects of school subjects in arranging the daily schedule. Interest and liking for a subject are more important than possible fatigue effects.
7. Nervous breakdowns are not caused by mental fatigue. The emotional factors in an individual's life are the causes. Worry about study, rather than too much study, is the potent factor.
8. Feelings of dissatisfaction or tiredness are not a good measure of efficiency.
9. Loss of efficiency in work may be postponed for a long time by intense motivation.
10. Industrial fatigue, as measured by output of work, includes mental and physical fatigue as well as the dissatisfactions and annoyances caused by continuous application to a task.
11. One or two night's loss of sleep does not seem to have any marked effect on mental tasks.
12. Drugs vary enormously with reference to their effects on mental work. Individual differences are marked here; the generally competent individual seems able to resist longer than the less generally competent.
13. Alcohol, even in small doses, acts as a depressant rather than as a stimulant.
14. Coffee in the customary small doses acts as a stimulant.
15. Tobacco seems to have very minor effects on habitual smokers.

Review

True-False Statements

As a review mark the following statements true or false:

1. Experiments show that children do best on arithmetic during the first period in the morning.
2. Mental "breakdowns" are generally caused by worry rather than by excessive mental work.
3. After four hours of mental work it is impossible to continue without serious loss of efficiency.
4. The effect of smoking on smokers in mental tests is always unfavorable.
5. The curve of satisfyingness always follows the curve of achievement.
6. The mental efficiency of school children on mental tests seems to be fairly constant during the school day.
7. Two nights' loss of sleep will not affect one's efficiency, if there is strong motivation.
8. Mental fatigue increases rapidly in the course of the usual school day.
9. We never find increased efficiency after two or three hours continuous work.
10. All work curves show an initial spurt.
11. Small doses of alcohol are stimulating and lead to increased scores on psychological tests.

References

1. Hollingworth, H. L.: "The Influence of Alcohol," *Journal of Abnormal and Social Psychology,* 1923–24, 18, 204–237 and 311–333.
2. Hull, C. L.: "The Influence of Tobacco Smoking on Mental and Motor Efficiency," *Psychological Monographs*, Volume 33, Number 3, 1924.
3. Knight, F. B., and Remmers, H. H.: "Fluctuations in Mental Production when Motivation is the Main Variable," *Journal of Applied Psychology*, 1923, 7, 209–223.
4. Laslett, H. H.: "An Experiment on the Effects of Loss of Sleep," *Journal of Experimental Psychology*, 1924, 7, 45–58.
5. Robinson, R. B., and Robinson, F. R.: "Effects of Loss of Sleep," *Journal of Experimental Psychology,* 1922, 5, 93–100.

CHAPTER XIV

MEASURING IMPROVEMENT IN SCHOOL SUBJECTS

School Improvement. — The improvement that the school is most interested in is that which it can bring about in its pupils. The school is an agency organized to bring about certain changes in the behavior of individuals. It is one of the many institutions trying to modify human behavior. Society determines what modifications the school should endeavor to make. The modifications thought desirable change from time to time and from nation to nation, as the history of education shows. The Greeks thought improvement in music and gymnastics was important; in the middle ages the schools considered improvement in Latin of first importance. None of these improvements are considered extremely vital in our schools to-day. Educational psychology attempts to study the changes or modifications which the school is trying to make in its pupils, whatever those modifications may be.

So far our study of modification or learning has been general, in an attempt to understand the universal principles at the basis of all learning. In this chapter we will be concerned with the specific modifications which the school is trying to make. The changes in pupil behavior which the school is trying to make are generally thought of as subjects which the pupil is learning. The modern school is, however, not only interested in teaching certain subjects in the narrow sense, but also in modifying the behavior of the child in regard to all those things summed up under such a phrase as "desirable citizenship." The school is attempting to produce individuals who will know certain things, who will act in moral and desirable ways, who will have certain appreciations and so on. Much of this is necessarily vague and particularly so when we descend from the abstract generalization to the concrete particular instance. Nevertheless, as the school is organized to-day, we see it very specifically trying to teach certain things, and so the psychologist has started to measure those specific things, hoping to move along to the measurement of the more intangible effects of schooling later on.

We really should not call them intangible, because any change made in a child will show itself in a change in his reactions. Tests can be constructed to measure whatever changes the school determines to make in a child. At the present time schools emphasize the acquirement of certain knowledges and skills. Hence the psychologist has set about the measurement of such knowledges and skills.

Achievement Tests. — In this way there have arisen many tests for school achievement. They are measures of the modifications brought about by the school. The score of any one such test represents for any given child a point on a learning curve for that particular function. If the tests are repeated over a period of time, the scores could be used to make learning curves for the function in question. Educational tests measure the improvement that the school has made up to a given time. They show what has been accomplished with the methods and materials used. They cannot show what might be accomplished with better methods or better management. They give a picture of the existing state of affairs. They have arisen and are in widespread use to-day, because they are accurate measures and because they are standardized measures. They are more accurate than the ordinary school examination. They are free from the personal bias of the teacher and their results can be compared with the results obtained in other schools all over the country.

The point we wish to emphasize here, however, is the important fact that all such tests are attempts to measure the modifications imposed upon original nature. They answer the question, "How much change has been effected?" Educational tests repeated from time to time give us a modification curve or learning curve for a particular individual. They show the rate and degree of modification taking place. Norms are the average amounts of modification achieved under present school conditions for children in general.

Learning Curves of School Subjects. — If we test a child again and again in a school subject, we will get a learning curve for that subject, just as we obtained learning curves for different functions, such as writing the alphabet backwards, typewriting and so forth in Chapter IX. Ordinarily exact records are not kept of a child's progress in any school subject. As the child moves from grade to grade the examinations become increasingly difficult and such changes in examinations make it impossible to construct a learning curve. However, by the use of standard educational tests, we may measure the progress of any child.

Figure 73 shows the progress curve in reading for a child who had been given the Stanford Achievement Test seven times at irregular periods during four years. The units of measurement here used are the educational age norms for the test. We notice an almost continuous increase in reading ability from July, 1923, to June, 1927. On the same figure there is also given the curve for arithmetic. The improvement here is steady from July, 1923, to June, 1926. There is a slight drop in the last test given in June, 1927.

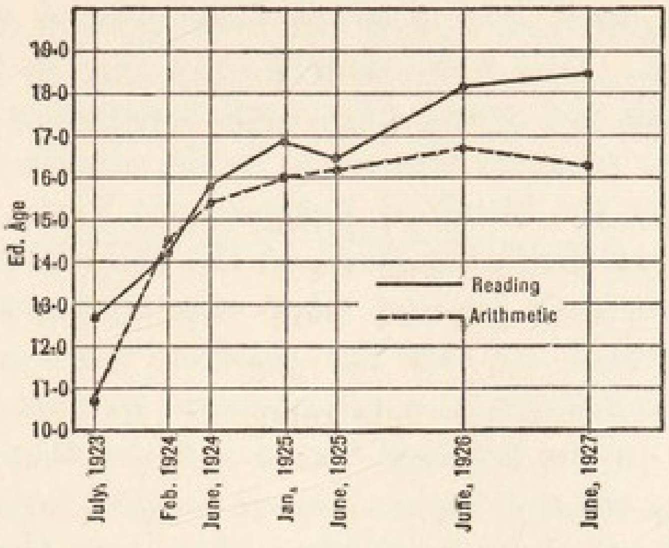


FIGURE 73. — Improvement in Reading and Arithmetic for one child.

Figure 74 shows two educational achievement curves for two other children tested by the Stanford Achieve-

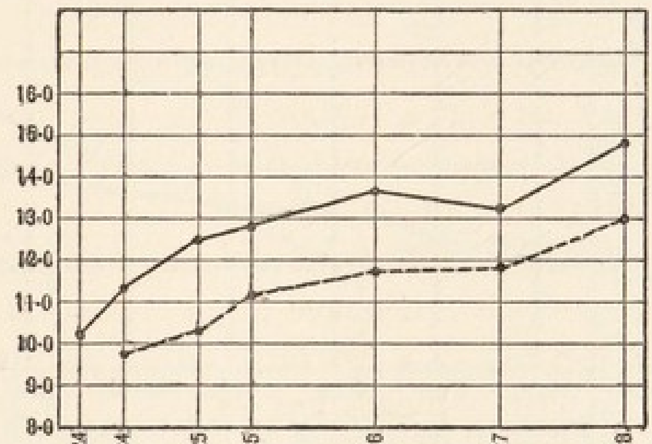


FIGURE 74. — Improvement in general educational achievement for two children.

ment Tests over a period of four and a half years. Again we note a fairly continuous and steady improvement with the exception of one test. It is interesting to note in the case of these three children the tendency to slow up or drop back in the June, 1927, test. Up to June, 1926 these three children had a very individual type of instruction which took into account their special needs and peculiarities. This was followed by a year in the public schools with the usual classroom instruction. During this year no progress was made in the standard subjects measured by the Stanford Achievement Tests. Perhaps the individual attention given to the children had raised their achievement beyond what was customary in the ordinary school, or else the children were engaged in other improvements and adjustments so that their improvement in the subjects tested suffered thereby. The second year of public school shows decided improvement. Repeated tests given to children show how they are improving or learning, and we may note that such curves are learning curves just the same as those discussed in Chapter IX.

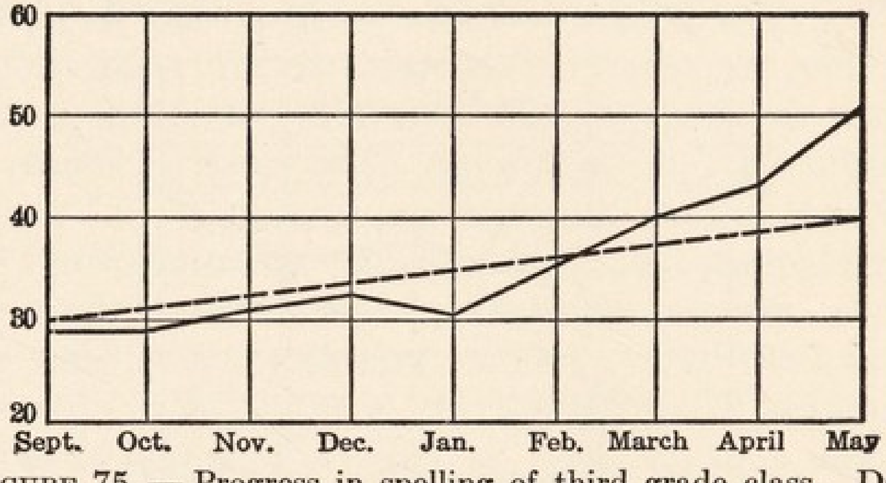


FIGURE 75. — Progress in spelling of third grade class. Dotted line represents standard scores. (From Starch's *Educational Psychology*.)

Composite Learning Curves. — Just as we have drawn learning curves of school subjects for individual children, so we can construct learning curves for groups of children. Figure 75 shows the learning curve in spelling, as measured by the Starch Spelling Test, for a third grade class. The class was tested each month during the school year and the solid line represents the improvement made, and we note an irregular but steady improvement from

September till May. The broken line is a hypothetical curve representing the probable improvement during the school year, as determined by the norms for the test.

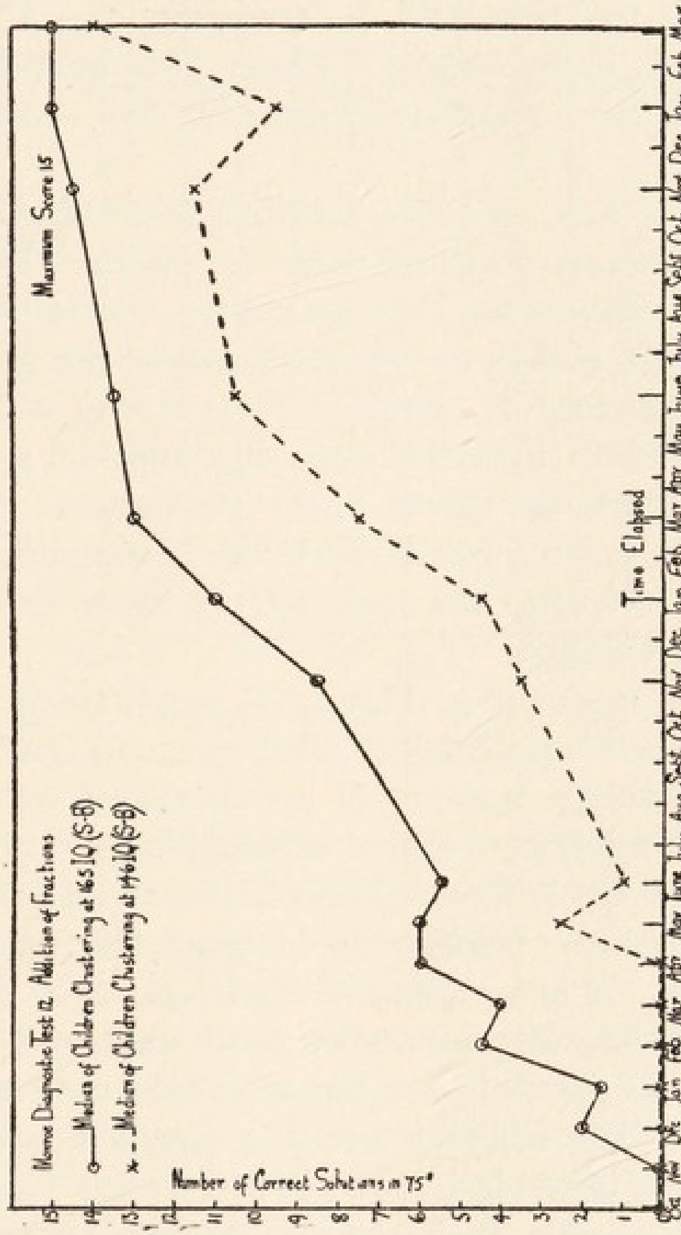


FIGURE 76. — Learning curves for the addition of fractions for two groups of superior children. (From Hollingworth and Cobb. *The 27th Yearbook of the National Society for the Study of Education.*)

Figure 76 is an interesting composite curve taken from Hollingworth's (1) investigations of superior children. It shows the improvement in the ability to add fractions, as tested by a specific test, for two groups of superior children tested many times during a period of three years. The solid line represents the progress of a very superior group with I.Q.'s around 165 and the broken line that of a superior group with I.Q.'s around 146. Notice how the very superior group keeps well ahead of the superior group, until the former approaches the maximum score on the test, and is, therefore, not able to show increased improvement. These two curves are learning curves for a specific function for two groups of children.

The norms of any well standardized educational test will show the average learning curve for children in general. If we use age norms, then we get the average learning curve by age, and if we use grade norms we get the average learning curve by grade. These curves will represent what children of various ages or grades are accomplishing in the average school under present-day conditions. They do not represent maximum accomplishment under ideal conditions, but they simply show us what children are now doing.

Such a curve is shown in Figure 77, constructed from the grade norms of the Stanford Achievement Test. It shows us the average number of points scored for each grade for the composite of school subjects included in the Stanford Achievement Test. Each half year of instruction shows a steady increase in learning the ordinary school subjects. It is to be noted that this curve differs from the other learning curves we have discussed in as much as it is not derived from repeated measures on the same children. The children tested in Grade V are not the same as those tested in Grade IV, and so on. To make the curve comparable with our other learning curves, the same group of children should have been tested each half year as they progressed through the grades. The curve in Figure 77 takes no account of the children who fall behind or drop out owing to lack of intelligence. Nevertheless, we may regard it as a general learning curve of the ordinary school subjects in the grades.

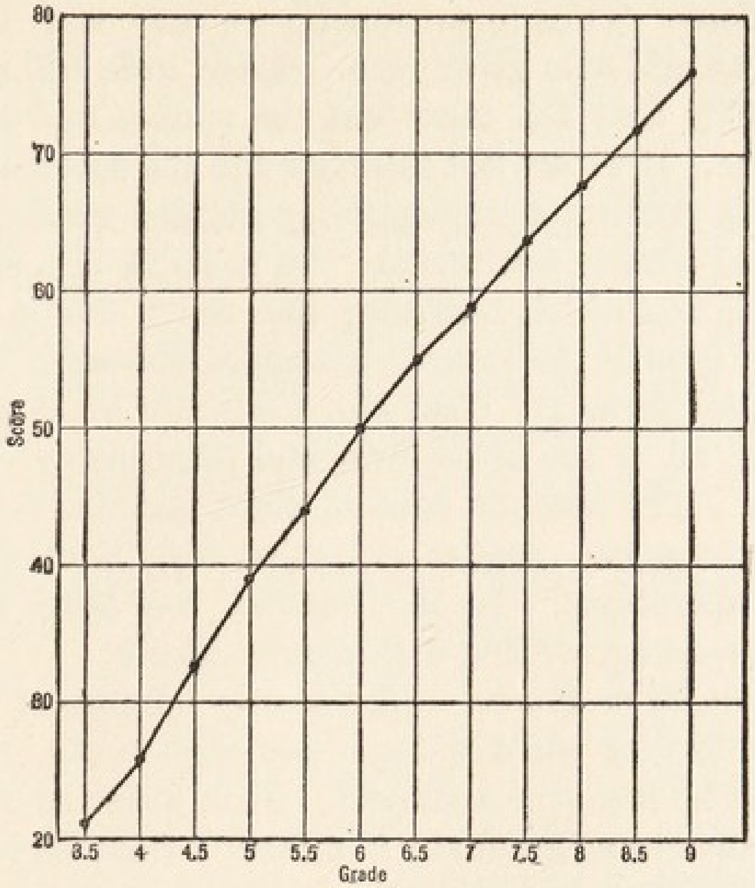


FIGURE 77. — Improvement in educational achievement from grade to grade. Stanford Achievement Test grade norms.

The Types of Tests. — Let us now make a brief survey of the numerous educational or achievement tests which have been constructed in order to help in the measurement of improvement in school work. Any attempt to describe these tests must take into account the many different aspects of a subject that can be measured. No single classification of educational tests is possible. They can, however, be roughly arranged (A) according to their purpose, or the aspect of any particular subject with which they are concerned, and (B) according to the field of knowledge or skill, that is the subject-matter, which they measure.

A. *Aspect or Purpose:* 1. *Amount, rate, or speed.* — A test may emphasize the amount accomplished in a given period of time, or the rate or speed at which certain things can be done. Thus in arithmetic we may wish to teach children to add at a given rate. Speed tests will measure how quickly they can carry out the various operations in arithmetic. If we are not interested in the speed at which our pupils perform arithmetical operations we should not give speed tests in arithmetic. An example of a standard arithmetic test which measures how much can be done in a certain time is the Courtis Standard Research Tests in Arithmetic, Series B. Test No. 1 consists of 24 addition examples, all of the same form and presumably of equal difficulty. The amount done in eight minutes forms the basis of the score. And so with other tests in other arithmetical operations. In like manner we have tests of speed in reading, writing and other subjects.

2. *Difficulty or Power.* — A test may emphasize the degree of difficulty which a pupil can master, the power or command he has over a subject. Such tests are made up of items increasing gradually in difficulty from very easy to very hard. The best "difficulty" or "power" tests arrange the items in steps of equal difficulty from easy to hard.

An example of such a test in Arithmetic is the Woody Arithmetic Scales. The addition scale, for example, is constructed of 38 addition examples, beginning with such easy items as $\begin{array}{r}2\\3\\\hline\end{array}$ and increasing gradually in difficulty through items like $\frac{3}{4}+\frac{1}{4}=$ to hard items like 25.091+100.4+25+98.28+19.3614= . How far the pupil can go up this ladder of difficulty will measure his power in addition.

In Reading, the Thorndike-McCall Reading Scales give a measure of difficulty, and so there are other tests in other subjects.

Difficulty or power tests are most useful in diagnosing the particular difficulties or mistakes made by a pupil. When they are used for this purpose they are called "diagnostic" tests. Thus the Charters Diagnostic Language Tests are specially constructed to diagnose the language difficulties of the pupils. By means of them we can find out or "diagnose" the language troubles of a pupil.

3. *Quality.* — A test may try to measure how well a person can do a certain thing, regardless of speed or of the difficulty of the operation. Such tests emphasize quality. The best example of this type is the Thorndike Handwriting Scale. It consists of a series of examples arranged in order from poor quality to good quality and the specific example to be rated is compared with the samples on the scale.

Quality again is important in English composition and there are several scales for its measurement.

These three aspects are perhaps the most important. There may be others. We might, for example, consider range as an aspect and construct a test to see how wide a range of information has been attained in such subjects as history or geography. And so there undoubtedly are other aspects. Any aspect which the school may from time to time emphasize will undoubtedly be measured by appropriate tests. Aspects overlap and most tests measure more than one aspect at the same time, but generally they emphasize one aspect in particular.

This brief treatment of the aspect of a subject measured by a test is important. We must always keep in mind precisely what we are measuring. There is no one test for Arithmetic, nor for Reading, nor for any school subject. We must measure one aspect at a time, and be careful to remember what that aspect is when we come to interpret our results. Furthermore we must not be so stupid as to blame the test for not measuring what it makes no claim to measure.

B. *Subject-Matter.* — Tests classified according to the field of knowledge or skill tested can be divided into two large groups: 1. General or composite tests including several fields of knowledge; 2. Special tests dealing only with one field of knowledge or with a limited portion of that field.

1. *General or Composite.* — Tests of this type include a number of school subjects and it is possible by their use to arrive at a general rating of the educational achievement of the child. A good example of such a composite educational test for the elementary school is the Stanford Achievement Test. It includes sub-tests in the following school subjects: — three different kinds of reading, two kinds of arithmetic, nature study and science, history and literature, language usage, and spelling. The total score on this test is a good measure of a child's general school accomplishment but this score tells us nothing about his ability in any specific subject. The total score of any given child can be compared to the norms and we can then tell how the child in question compares with other children of his age or grade. Similarly the average scores for classes or schools can be compared with each other and with the general standards.

The Stanford Achievement Test is a fairly long and thorough test and because of this it gives fairly good independent measures of the several subjects entering into the composite. So in addition to the general educational standing of the child, we can obtain his standing on particular subjects, such as reading, arithmetic and so on.

An example of a much briefer test covering the ordinary subjects in the elementary field is the Pintner Educational Survey Test. This also contains tests in reading, arithmetic, history, geography, language and so on. The tests are very brief and little reliance can be placed on the rating of any specific subject. Nevertheless, the total score on the whole test gives a fair measure of a child's general educational standing. Tests of this type are generally called survey tests, because they are designed for a general survey of a school or school system, and they are not meant to give detailed information about specific children.

At the high school level, we have the Iowa High School Content Examination which consists of four sub-tests as follows: (1) English, Literature, etc.; (2) Mathematics; (3) Science; (4) History and Social Science. It is a thorough examination of the usual things taught in high school. A general score for the whole test may be obtained as well as separate scores for the four separate parts of the examination.

2. *Special Tests.* — The number of special tests for separate subjects is very great and it is not our purpose to attempt to describe them all. Still less is it possible to evaluate them or contrast the merits of one with those of another. Indeed the point of view we shall emphasize is that most of them possess merit of one kind or another, and the important thing for the teacher to remember is the particular aspect of a subject that each one is measuring. No one test will measure all the objectives of arithmetic teaching in the school. We must carefully choose the test that best fits the aim or objective we want to measure.

We shall run over some of the more important subjects in the curriculum and make comments on some of the tests in each field.

a. *Arithmetic.* — Arithmetic is well represented by measuring instruments, because it is relatively easy to construct tests in this field. The content is fairly uniform in all schools and the scoring of such tests is definite and easy. If we want to measure speed and accuracy in the four fundamental operations we can use the Courtis Standard Research Tests, Series B. These tests will not tell us how hard or complicated examples in addition or subtraction our children can solve, but merely how quickly and how accurately they can add, subtract, multiply or divide arithmetical examples of a given degree of difficulty. They will merely answer that question and they should not be used for any other purpose.

If, however, we are interested in the problem of difficulty and wish to know how high on the scale of difficulty or complexity a child can rise, then we will use the Woody Arithmetic Scales or some similar type of measuring instrument. The Woody tests are difficulty or "power" tests in the four fundamental operations.

Difficulty or "power" in general in the four fundamentals can be measured by the Woody-McCall Mixed Fundamentals, which as the name implies is the same type of test as the Woody test except that here the four fundamental operations are mixed together.

In addition to the fundamentals of arithmetic, problem solving is very important, and none of the scales dealing with the fundamentals will give us a measure of this. Therefore, if we wish to measure problem solving rather than dexterity in the fundamentals we must make use of some such test as the Buckingham Scale for Problems in Arithmetic. There are three scales for the different grade levels from III to VIII, and the use of this test will give us a good measure of the child's ability to solve arithmetical problems as distinct from his ability in the four fundamental operations. The diversity of tests in arithmetic demonstrates very well the necessity for carefully choosing the appropriate test for that specific thing which we wish to measure.

As we go from the elementary to the high school, the problem becomes still more complex, for here the field widens into algebra and geometry. The Hotz First-Year Algebra Scales cover the main operations in elementary algebra.

The Columbia Research Bureau Plane Geometry Test is a very thorough test in geometry for high school students. There are, in addition, several other tests in algebra and geometry, and each has its particular merit.

b. *Reading.* — When we pass from arithmetic to reading, we pass from a relatively simple to a relatively complex school subject. For reading comprises such abilities as being able to articulate printed words placed before one, being able to perceive silently the symbols, remember the general content, respond to questions about it, understand the meaning of isolated words, carry out printed directions and so forth. The functions measured by "reading" tests may therefore, be basically very diverse. Ordinarily reading tests are divided into three groups, silent reading, oral reading and vocabulary tests.

Silent reading tests are numerous. They consist of passages to be read followed by questions to be answered in order to test the comprehension of the passages. The Thorndike-McCall Reading Scales for Grades III to XII are excellent examples of tests for the measurement of comprehension in reading. For example, the pupil is asked to read a short passage and then, with the passage still before him, answer some questions:

"On Monday Dick saw a red fox, a gray squirrel and a black snake in the woods. The next day he saw a brown rabbit and five brown mice in the field. He killed the fox and all the mice, but let the others live.

"What did Dick do to the fox? ....
"What color was the rabbit? ....
"Where did Dick see the fox? ...."

The tests are primarily power tests in the sense that they measure ability to comprehend. They do not measure rapidity of reading. They have a high correlation with general intelligence tests of the verbal type, partly because the latter assume the ability to read with understanding. The Burgess Picture Supplement Scale, designed for Grades III to VIII, consists of passages containing directions to be carried out. These directions refer to some picture and the child responds by making some mark or drawing on the picture. The child does not have to respond by a word, phrase or sentence as in the Thorndike-McCall Tests. The Monroe Standardized Silent Reading Tests offer passages for reading with objective type questions to measure comprehension. They give measures both of comprehension and of rate of reading.

Oral reading tests measure the ability to read orally without error. Errors consist of mispronunciations, omissions, substitutions, insertions and repetitions. The Gray Oral Reading Tests cover this field very well. Measurement of oral reading must be done individually and it is therefore very time-consuming. Again, since oral reading is not very important in daily life, the need for such tests is limited.

Vocabulary tests confine themselves to a measurement of the knowledge of words. They generally have a high correlation with verbal intelligence tests, many of which contain vocabulary tests as sub-tests of the composite test. Vocabulary tests are generally constructed on the objective test pattern so that they are simple to give and easy to score. Knowledge of words is undoubtedly one of the most important factors in reading ability. The Thorndike Test of Word Knowledge is a good sample of this type of test. The test consists of one hundred graded words, each of which has along with it five other words, and the child has to mark the word which means the same or nearly the same thing. Here are samples from the list of 100:

No. 1. afraid: full of fear — possible — necessary — raid — ill.
No. 31. anon: year — sometimes — hitherto — again — now.
No. 54. ardor: anger — need — zeal — difficulty — pagan.
No. 97. madrigal: song — mountebank — lunatic — ribald — sycophant.

Reading has so many aspects that no one test can measure all of them. The most thorough-going attempt to construct tests for many of these aspects is seen in the work of Gates. He has constructed special tests for the primary grades; silent reading tests for Grades III to VIII of four different types: Type A, to appreciate the general significance of a paragraph; Type B, to predict the outcome of given events; Type C, to understand precise directions; Type D, to note details; a graded word pronunciation test; and a phonetic ability test. These tests of Gates cover the field of reading very thoroughly. They are absolutely indispensable for diagnosing the reading difficulties of pupils. Each test has its special function. This whole battery of tests illustrates excellently the necessity for deciding definitely just what we want to measure before we start our testing. We cannot just measure "reading" in general. We must decide specifically what aspect of reading we are going to deal with.

c. *Handwriting.* — The essential factor in handwriting is quality or legibility. A certain degree of legibility is necessary as a minimum and a large amount of it is highly desirable. Here we are dealing then with the measurement of a product quite unlike reading or arithmetic. The scales constructed for this purpose consist of a series of sample products ranging from bad to good by gradual steps. The product to be rated is compared with the samples of the scale and given the value of the sample to which it most nearly approximates. The two most useful handwriting scales are the Thorndike Scale and the Ayres Scale. As can be readily imagined from our description of these scales, the measurement of handwriting cannot be as objective as the measurement of arithmetic or reading. Raters may vary as to their judgment of the quality of any given sample of handwriting. A certain amount of practice is necessary in order to use these scales properly. Care must be taken in comparing handwriting scores obtained by different raters. Nevertheless, in spite of all these difficulties, inherent in any quality scale, scales for the measurement of handwriting are of great value.

d. *Spelling.* — Ability to spell words according to the existing convention of the day is, strangely enough, very highly rated by educators and therefore many scales or tests have been constructed for the measurement of this function. These are generally based upon commonly used words and are arranged in order of difficulty. The Morrison-McCall Spelling Scales and the Iowa Spelling Scales are useful in the elementary grades, while the Brigg's Scales are of value in the high school.

e. *English.* — Under this heading we have to consider such diverse subjects as Composition, Language, Grammar and Literature. Composition, the ability to express oneself properly in writing, is one of the most important objectives in education. What is measured here is the product produced by the child, and, as in Handwriting, this is done by comparison with a standard scale of samples arranged in order of merit. The Thorndike Extension of the Hillegas Scale contains fifteen degrees of merit from very poor to very good compositions. What is judged by the rater is the general merit of the composition and this general merit is compared with the standard samples of the scale. The Van Wagenen English Composition Scales consist of separate scales for exposition, narration and description. Each composition is rated with reference to thought content, structure and mechanics. It is an attempt to analyze more specifically the elements that enter into composition ability. The rating of composition, like the use of any quality scale, is difficult and what has been said above with reference to the use of handwriting scales applies here with still greater force, because of the complexity of the thing we are rating. Ordinarily a fair amount of practice with any scale is necessary before reliable judgments are secured. Even with the use of the scales, individual teachers may differ greatly in their judgments. Nevertheless, the use of any such scale by a teacher will tend to make his ratings more and more uniform, and for educational research purposes the use of a scale by trained workers is indispensable.

Language tests measure the ability to use correct English and are of course more objective and easier to use than are the composition scales. The Wilson Language Error Test presents a short story and the pupil has to correct all the mistakes in the story. The Charters Diagnostic Language Tests expect the child to detect and correct errors. The Kirby Grammar Test, in addition, calls for the choice of the grammatical principle concerned.

Attempts to measure what is taught under the heading of English literature in schools are very diverse and very limited. Probably this is owing to the very hazy and disputed objectives of literature teaching in the schools. There are such tests as the Abbott-Trabue Exercises in Judging Poetry, and the Van Wagenen Reading Scales in English Literature. Literature information can be tested readily by the ordinary information type of test.

f. *Geography.* — Here the problem is largely one of testing geographical information. The Gregory-Spencer Geography Tests are a good example. They cover such phases of geography as trade routes, causal geography, description and map location, physical and commercial geography. The Posey-Van Wagenen Geography Scales attempt to cover a wider range, and include a Thought Scale which is supposed to measure a child's ability to think about geographical problems.

g. *History.* — Here, as in Geography, the field is well represented by tests. The Barr Diagnostic Tests in American History measure separately comprehension of material read, chronological judgment, judgment of evidence, evaluation of the importance of facts, and the understanding of cause and effect relationships. Other tests stress historical information. The Van Wagenen American History Scales measure information, thought and character judgment.

h. *Foreign Languages.* — There are now many good tests for the measurement of ability in French, Spanish, Latin and German. The Columbia Research Bureau Tests in French, Spanish and German are good examples of modern work in this field.

i. *Science.* — The large field covered under this general heading leads to many diverse tests. There are general science tests, such as the Ruch-Popenoe General Science Test containing items from botany, chemistry, physics, zoology, astronomy, geology and so forth. Then there are tests in the separate sciences, such as the Powers General Chemistry Test, the Ruch-Cossmann Biology Test and the Iowa Physics Test. Naturally in this wide field there is room for a great many tests.

j. *Miscellaneous Subjects.* — In addition to the more or less standard subjects which we have surveyed above, there remain such subjects as music, hygiene, cooking, sewing, shop work and in the high schools the vocational subjects, such as typewriting, stenography and the like. There are educational achievement tests or scales in practically all of these subjects. In some the measurement is rather meager, but in others the tests are well-established. In all the fields the possibility of measurement has been demonstrated.

More elaborate accounts of achievement tests will be found in the works of Kelley (2), Monroe (3), Symonds (5), and Ruch and Stoddard (4).

Tests and School Objectives. — Our survey of the educational tests and scales shows very definitely that anything which the school sets out to teach can be measured by suitable tests. Training, schooling, education or whatever else we call it means, from the psychological standpoint, the modification in some way or other of the reaction habits of the child. These reaction habits lend themselves to measurement and our educational tests are the measures of the different kinds of reactions which the school endeavors to bring about in the child.

Frequently the aim of education is expressed in lofty and general terms, such as better citizenship, a higher morality, a Christian life and so on. So long as these terms remain undefined and vague, there is considerable agreement among people as to the desirability of these objectives. But, as soon as we attempt to be more specific as to just what is good citizenship, as to just what is the Christian life, disagreement at once begins. We cannot set out vaguely to make children good citizens. We have to be concrete and definite, and set out to teach them so much reading and writing and history and geography and habits of health and definite ways of reacting to this, that, and the other thing in their environment. Immediately we do this, we arrive at tangible reactions which are open to psychological measurement. Let the educator set out definitely to modify the reactions of a child in a particular way, and the psychologist will be able to measure the amount of change or modification that is taking place.

Interpretation of Educational Ratings. — In all of the educational achievement tests discussed above there results from the testing some kind of score or numerical value. These scores vary greatly from test to test and in order to be able to interpret them, they must be transmuted into some standard rating.

One of the best ways of interpreting any particular score is to compare it with the distribution of a large and random sampling of other children's scores. Since most, if not all, distributions of educational scores tend towards a normal distribution, the usual way is to compare the individual score with the standard deviation of the comparable group. In this way a score is expressed as so many times the standard deviation of the standard group. Standard deviations are calculated from the mean and therefore we have plus and minus quantities. These are difficult to handle. Because of this, modifications of the direct standard deviation units are more customarily employed. The best known and most useful modification is the T-Scale devised by McCall. A T-unit is one-tenth of the standard deviation of a random sampling of twelve-year-old children. T-units run from zero to one hundred, and fifty is the mean score for twelve-year-olds. The distribution of any age group can then be interpreted in terms of the standard twelve-year-old group. It is also possible to calculate T-scores for every age group or indeed for any group of children. Tables of T-scores are frequently given for standard educational tests.

Another method of interpreting educational test scores is to convert them into educational ages. Just as an intelligence test score is converted into a mental age (M.A.), so an educational test score is converted into an educational age (E.A.). The E.A. is obtained by an age standardization for those ages for which the test is most useful. The meaning of the E.A. is simple and easily grasped. An E.A. of 10 on a test means that the child in question is achieving what ten-year-olds ordinarily achieve; an E.A. of 12 means he is up to average twelve-year-old performance and so on. The scheme is simple and useful but not as accurate as any of the standard deviation methods. The drawback of the E.A. method is that the E.A. becomes less and less reliable as we proceed to the higher ages. In these higher ages an E.A. becomes the average score of a much more selected group than the E.A. of the lower ages. Whereas an E.A. of 11 or 12 derived from an extensive standardization of elementary school pupils may be said to represent fairly the achievement of twelve-year-olds in general, an E.A. of

15 or 16 obtained from school children cannot in any sense be considered the average achievement of fifteen- or sixteen-year-olds in general. These fifteen- or sixteen-year-olds are a very selected group. The E.A. should not, therefore, be used in the high school, and teachers should be aware of the fictitious nature of E.A.'s in the upper ages when they make use of them to interpret very high achievement scores on any test.

If we use the E.A. for any educational achievement test, we have the Educational Age on that particular test. Sometimes we are making use of several tests in different subjects, and so we obtain several E.A.'s. To distinguish between these it is quite customary to give them specific names, and so we hear of Arithmetic Age, Reading Age, Geography Age, and so on, and sometimes these are contracted into symbols such as A.A., R.A., and the like. They are all, however, educational ages of some sort or other.

Following the analogy of the Mental Age and Intelligence Quotient, we have for educational tests the Educational Quotient accompanying the Educational Age. The E.Q. is the E.A. divided by the C.A., just as the I.Q. is the M.A. divided by the C.A. The meaning of the E.Q. is, therefore, similar to that of the I.Q., keeping in mind the different sort of thing we are measuring. Thus an E.Q. of 100 is normal or average for the age in question; an E.Q. of 125 is above average and an E.Q. of 80 is below. The E.Q. of 80 tells us that for some reason or other the child in question is doing quite inferior work for his age. The reason may be lack of intelligence, faulty grade placement, faulty teaching and so forth. The E.Q. cannot reveal the reason. The E.Q. of 125 tells us that the child is doing much better educational work than is usually accomplished by children of his age. This may be due to superior intelligence, superior schooling, exceptional opportunity and so forth. We must remember that the E.Q. is in terms of the average educational accomplishment of children of different chronological ages. Its reference point is always the accomplishment of the average child under present-day instruction, with average opportunities. It does not refer to what might be accomplished under ideal conditions.

All educational tests have furthermore another type of rating which can be used in addition to either of the above methods. This rating is a grade rating of some type or other. Such ratings are based upon grade standardizations, in contrast to the age standardizations discussed above. The grade ratings tell us how any particular child or class compares with the average for grades in general. The grade rating is important particularly from the supervisor's or superintendent's point of view. Its significance, however, can only be determined by taking into consideration such factors as the intelligence of the children making up any particular grade, and the chronological age of the children to determine whether they are under- or over-age for the grade in question. We shall see how these factors are important in our discussion of the uses of educational tests.

Summary

1. This chapter deals with the measurements of the modifications which the school is trying to make in its children.
2. An educational test score represents a point on a learning curve for some child in some particular function.
3. Repeated tests allow us to plot learning curves for particular subjects.
4. Progress in educational subjects is not necessarily steady and continuous.
5. The standard grade or age norms for a test may be considered as composite learning curves for children in general, assuming such standards to be based upon the same kind of selection of children at each age or grade.

6. Educational tests may be considered according to (1) their general purpose and (2) the subject matter measured.
7. Purposes may be to measure speed or power or quality.
8. Tests considered from the aspect of subject matter may be further divided into: general or survey tests covering several fields of knowledge; and specific tests dealing with one field only.
9. Examples of these general and special tests have been given and it is seen that almost every subject of school instruction is now covered by tests. This applies to both the elementary and the high school. Some fields have been well covered and others very meagerly.
10. It is concluded that any specific objective which the school may set up is measurable by means of achievement tests.
11. Scores obtained on educational tests must be transmuted in order to be interpreted.
12. Two common ways of transmuting scores are the standard deviation method and the educational age method.
13. There are limitations to the usefulness of Educational Age and Educational Quotient.
14. All educational tests give us the opportunity of interpreting the scores in terms of grade norms.

Review

True-False Statements

As a review mark the following statements true or false:

1. Achievement tests should determine the ideal type of subject matter to be taught in schools.
2. Achievement test scores are generally more reliable than teachers' marks.
3. Repeated educational tests of the same child give us a learning curve for that child.
4. No one achievement test will measure all the various aspects of any school subject.
5. Most achievement tests are merely speed tests.
6. Which arithmetic test we choose will depend upon what specific aspect of arithmetic we desire to measure.
7. Power tests are those which emphasize how quickly a child can do the test.
8. The educational quotient is derived in the same way as is the intelligence quotient.
9. Complete this sentence:
   The E.Q. is the quotient resulting from dividing the ——— by the ———.

Advised Reading

Gates, A. I.: *The Improvement of Reading*, New York, 1927.

References

1. Hollingworth, L. S., and Cobb, M. V.: *Twenty-Seventh Yearbook of the National Society for the Study of Education*, Part II, 1928, Chapter 1.
2. Kelley, T. L.: *Interpretation of Educational Measurements*, Yonkers, 1927.
3. Monroe, W. S., et al.: *Educational Tests and Measurements*, Boston, 1924.
4. Ruch, G. M., and Stoddard, G. D.: *Tests and Measurements in High School Instruction*, Yonkers, 1927.
5. Symonds, P. M.: *Measurement in Secondary Education*, New York, 1927.

List of Tests Referred to in this Chapter

Composite Educational Tests

Stanford Achievement Test: World Book Co., Yonkers.

Pintner Educational Survey Test: College Book Co., Columbus, Ohio.

Iowa High School Content Examination: Bureau of Educational Service, University of Iowa.

Arithmetic and Mathematics

Courtis Standard Research Tests, Series B.: Courtis, Detroit, Michigan.

Woody Arithmetic Scale: T. C. Bureau of Publications.

Woody-McCall Mixed Fundamentals: T. C. Bureau of Publications.

Buckingham Scale for Problems: Pub. Sch. Publishing Co.

Hotz Algebra Scales: T. C. Bureau of Publications.

Columbia Research Bureau Geometry Test: World Book Co., Yonkers.

Reading

Thorndike-McCall Reading Scales: T. C. Bureau of Publications.

Burgess Picture Supplement Scale: Russell Sage Foundation.

Monroe Silent Reading Tests: Pub. Sch. Publishing Co.

Gray Oral Reading Test: Pub. Sch. Publishing Co.

Thorndike Test of Word Knowledge: T. C. Bureau of Publications.

Gates Silent Reading Tests: T. C. Bureau of Publications.

Handwriting

Thorndike Handwriting Scale: T. C. Bureau of Publications.

Ayres Handwriting Scale: Russell Sage Foundation.

Spelling

Morrison-McCall Spelling Scale: World Book Co., Yonkers.

Iowa Spelling Scales: Pub. Sch. Publishing Co.

Briggs' Sixteen Spelling Scales: T. C. Bureau of Publications.

Composition

Thorndike Extension of Hillegas Scale: T. C. Bureau of Publications.

Van Wagenen English Composition Scales: World Book Co., Yonkers.

Language

Wilson Language Error Test: World Book Co., Yonkers.

Charters Diagnostic Language Test: Pub. Sch. Publishing Co.

Kirby Grammar Test: University of Iowa.

Literature

Abbott-Trabue Exercises in Judging Poetry: T. C. Bureau of Publications.

Van Wagenen Reading Scales in English Literature: Pub. Sch. Publishing Co.

Geography

Gregory-Spencer Geography Test: University of Cincinnati.

Posey-Van Wagenen Geography Scales: Pub. Sch. Publishing Co.

History

Barr Diagnostic Tests in American History: Pub. Sch. Publishing Co.

Van Wagenen American History Scales: T. C. Bureau of Publications.

Foreign Languages

Columbia Research Tests: World Book Co., Yonkers.

Science

Ruch-Popenoe General Science Test: World Book Co., Yonkers.

Powers Test for General Chemistry: World Book Co., Yonkers.

Ruch-Cossmann Biology Test: World Book Co., Yonkers.

Iowa Physics Test: Pub. Sch. Publishing Co.

CHAPTER XV

THE USES OF EDUCATIONAL TESTS

Mental Ability and Educational Achievement. — The intelligence test attempts to answer the question, "How much *can* this child achieve?" The educational test gives an answer to the question, "How much *has* he achieved?" The two instruments are measuring a child's reactions from different viewpoints. Sometimes indeed the very same type of reaction is used in the two types of tests, as when reading reactions and arithmetical reactions are included in intelligence tests. But usually the reactions called for in intelligence tests are more general. They are gathered from a wider field. Whereas the reactions called for in educational tests are more specific. They are gathered from a narrower field. They are limited very specifically to what the school is concerned with teaching. The intelligence test is essentially prognostic in function. It looks towards the future and predicts the type of achievement of which the child should be capable. It is this predictive aspect of intelligence tests that makes them so valuable. They indicate the amount of education a child should be capable of achieving. If a child has a high I.Q. his educational achievement should be high. If he has a low I.Q. we cannot expect his educational achievement to be up to the average for his age. An inherently bright boy should do better than one of merely normal ability, and it is no credit to the bright boy to achieve the average educational accomplishment for his age. To men of genius we look for works of genius. From men of feeble capacities we do not expect works of genius nor even average products.

The A.Q. Technique. — In this sense, therefore, all educational testing must be interpreted in the light of the intelligence of the children tested. To evaluate correctly the educational achievement of a pupil, we must have an accurate measure of his intelligence. In the actual measurement of pupils, several techniques have been developed to accomplish this. The easiest to understand and one frequently used is the A.Q. technique. This is built upon the I.Q. and the E.Q. A child of a given I.Q. is expected to attain a similar E.Q. This ratio of E.Q. to I.Q. is called the A.Q. or Achievement Quotient. The A.Q. is thus $\frac{\text{E.Q.}}{\text{I.Q.}}$ or $\frac{\text{E.A.}}{\text{M.A.}}$

If the E.Q.'s and the I.Q.'s are very similar we shall have A.Q.'s clustering around 100, and this means that educational attainment is keeping step with intellectual ability. If the E.Q. is appreciably below the I.Q., then our A.Q.'s will be small or appreciably below 100. This means that all the intelligence of the child is not being used in his educational work. For some reason or other he is not using all his intellectual capacity. The reason for this, of course, will not be given by our measurements. We must look for the reason elsewhere. A very common one is faulty grade placement of children of high I.Q. They are frequently in grades doing relatively simple work and hence they do not have the opportunity to accomplish more on their educational tests. Other reasons may be laziness, lack of proper motivation, illness and so on. If the E.Q. is appreciably above the I.Q. then we obtain A.Q.'s above 100. These indicate children who are accomplishing more than is ordinarily accomplished by children of their I.Q. level. We are reminded here of what was pointed out in the previous chapter, namely, that E.A.'s and E.Q.'s are based upon normal or average accomplishment and not upon maximum accomplishment. Hence it is perfectly possible for a child at any given

I.Q. level to achieve more than is ordinarily accomplished. This is brought about frequently by pushing some children faster than usual, by stimulating them to work by placing them in a grade where the mental age level is higher than their own and trying to keep them up to standard. This condition is frequently found among docile, industrious, dull children. They advance with their class because they are of the right chronological age and because their docility and industry impresses the teacher. They are constantly given extra attention and training by the teacher because they are the slowest in the class. All of this tends to raise their E.A. above that of children of like M.A. and hence we get A.Q.'s above 100. This illustrates very well the fact that our educational test standards are norms of average accomplishment and not standards of maximum accomplishment under ideal conditions. The average A.Q. of a class will be a measure of the efficiency of the class and similarly with the average of a whole school or school system.

The F Score. — The standard deviation methods of rating educational tests can also be used to evaluate educational achievement in terms of mental ability. The difference between the sigma position of the pupil on the educational and the mental tests is such a rating. In McCall's T score technique we have the F or efficiency score which is the achievement T score plus 50, minus the intelligence T score. The plus 50 is merely included so that the average or normal shall be 50, as in the T scale.

Limitations of These Techniques. — In addition to these A.Q. and T and F techniques, there are several others, but these will suffice as the most commonly used methods. All of these techniques must be used cautiously. They all depend upon a very thorough standardization of the educational and intelligence tests used. They all assume that the standardization groups of the two sets of tests, educational and intelligence, are similar if not identical. Obviously, if this is not the case, it will be impossible to interpret the A.Q.'s or F's in the usual way. If the intelligence test has been standardized by means of a poor group of children and the educational test by means of an average group, then when we come to use these standards on an average group of children, our I.Q.'s will be increased above what they really should be, and our E.Q.'s will be according to expectation. Not knowing the faulty standardization of the intelligence test, there will be nothing to show us the distortion of the I.Q.'s. Using, then, these distorted or elevated I.Q.'s, we will obtain depressed A.Q.'s. These we may interpret as due to poor teaching or what not, instead of to faulty standardization of the measures used. Ideally these techniques should not be used except with intelligence and educational batteries standardized on the same group of children.

General Uses of Educational Tests. — We can classify some of the common uses of educational tests as follows:

1. Survey and Inventory Purposes.
2. Efficiency of Instruction.
3. Sectioning and Promoting.
4. Diagnosis of Individual Peculiarities.
5. Motivation.
6. Research.

1. *Survey and Inventory Purposes.* — No school survey nowadays is complete without the use of objective tests and scales for the measurement of the school product. All the paraphernalia of modern education, buildings, playgrounds, superintendents, teachers, books and equipment, are after all solely useful in so far as they are modifying the child's reactions in definite desired ways. The tests and scales are measures of such modifications, and hence they help to give an answer as to the degree in

which the school equipment and staff are justifying themselves in the product produced. The results of such tests may be used to compare city with city, or the results in any one city may be compared with the norms of a well standardized test to see if the city in question is reaching the standards of the country at large. In making such comparisons there is obvious need for taking into account the intelligence of the children, and also the chronological age of the child where grade norms are used.

The tests used in such surveys need not be long and detailed. If class averages are merely to be made use of, the composite brief survey test is to be selected because of its economy in operation, even though it may not be very reliable as far as individual scores are concerned. Sometimes a particular subject may be measured in detail and then of course a special test for that subject should be used.

The use of a test for comparing city with city and with the standard for the country at large is shown in Table 9, which is taken from a survey of the schools of Tampa,

TABLE 9

ACHIEVEMENT IN SILENT READING IN TAMPA, FLORIDA, SCHOOLS FOR WHITE CHILDREN COMPARED WITH OTHER PLACES
MAY, 1925

Thorndike-McCall Test for Understanding of Sentences

City	Grade			
	4 Senior	*6 Senior*	*7 Senior*	*8 Senior*
TAMPA, Florida	39.7	50.0	56.0	61.0
Baltimore, Maryland	43.3	53.1	58.1	61.2
33 Wisconsin Cities	40.9	52.6	55.3	58.0
Paterson, N. J.	35.5	49.0	51.7	53.5
St. Paul, Minn.	41.8	53.5	58.0	62.5
Louisville, Ky.	39.1	51.7	59.8	60.7
Hackensack, N. J.	41.4	51.9	56.2	——
STANDARD	41.8	53.7	58.3	59.6

Florida (Strayer, 2). The table shows the average scores made by pupils in various grades in different cities. From this we see how Tampa compares with other cities, and by comparing the scores of Tampa with the standard scores on the bottom line, we note that Tampa in general falls a little behind the country at large.

In another survey (Port Arthur, Texas) standard tests are used to answer the following four questions:

"1. How well are the pupils in the schools of Port Arthur doing in each of the subjects for the grade they are in?
"2. How well are they doing for their ages?
"3. Are they achieving up to their ability to achieve?
"4. What are some of the apparent difficulties and short-comings?"

2. *Efficiency of Instruction.* — Educational tests are invaluable for supplementing and checking the judgments of supervisors and superintendents with reference to the efficiency of their teachers. To depend solely on the impressionistic method of dropping in to a classroom to judge the efficiency of a teacher means to be at the mercy of countless chance incidents, of numerous extraneous factors which warp our judgments. The total personality or some special mannerism of a teacher may prejudice the supervisor in one way or another. Hence the value of objective tests in addition to these personal judgments.

In the use of tests in this connection, it is obvious at once that intelligence tests must be used in order to evaluate the work of the teacher. A lower educational score in a class made up of inferior children may mean much better instruction than a higher score in another class of superior children. The difference between the average educational achievement score of two classes cannot directly measure the efficiency of instruction. We must interpret this in the light of the intelligence, the chronological age and the previous instruction of the children.

When this is done the educational achievement scores will give credit where credit is due. Many a good teacher has been blamed for poor teaching because the poor quality of the intelligence of her children has not been recognized. Many an indifferent teacher has been commended for good work simply because she has had a class of superior children who have made progress in spite of her.

3. *Sectioning and Promoting.* — In large schools where there are several sections of the same grade, educational tests along with intelligence tests, can be of great service in determining the children that should be placed in each section. The best instruction can be given to a group which is homogeneous mentally and educationally. A homogeneous intelligence group means a group of children of similar M.A.'s and I.Q.'s. Such a group is growing mentally at the same rate. They will be able to keep pace with each other in learning and if they all start from the same starting point in any subject, i.e. are homogeneous educationally, we have an ideal group for mass instruction. The proper use of tests will aid us to approach such an ideal.

Fast-moving and slow-moving sections can be more accurately determined by means of a judicious use of educational and intelligence tests. Some schools adopt the so-called three-track plan of fast, average and slow classes. There is no special virtue in two, three, four or any number of tracks. This may be determined by the number of pupils at any one grade level. The important thing is to have the sections as homogeneous, mentally and educationally, as is possible.

Educational tests may further help in determining promotion from one grade to another or from one section to another. They are merely to be considered a help in this matter, for they merely give a measure of a small sample of the work of the class. In the more individualized instruction of the freer type of school, educational tests may be used to check up on the work of the individual, to determine his progress, to decide whether he has mastered a certain unit of subject matter and should be allowed to go on to the next. Furthermore, in all schools standard educational tests are of the greatest help in the grade placement of new pupils. The probable grade levels of such new pupils may be very quickly determined.

The use of educational tests in sectioning, promoting and placing pupils gradually merges into their more general use in educational and vocational guidance. In this connection an educational achievement score is but one small item of the many that must be assembled in attempting to give sound educational or vocational advice.

4. *Diagnosis of Individual Peculiarities.* — Progress in any field means overcoming the difficulties that confront the individual. What is easy to some, is a stumbling-block to others. To help the individual, the teacher must know what are his particular difficulties. Educational tests help to discover the particular weaknesses of each child.

Almost any of the tests will give some kind of a diagnosis of the individual case, but obviously the survey type of test is too short to be of any practical use. In general the more specific and detailed a test is, the more useful will it be for diagnostic purposes.

The best example of achievement tests constructed for diagnostic purposes is to be found in the Gates Reading Tests (1). There are four types of tests:

Type A — Reading to Appreciate the General Significance of a Paragraph.
Type B — Reading to Predict the Outcome of Given Events.
Type C — Reading to Understand Precise Directions.
Type D — Reading to Note Details.

If we wish to know something of a child's reading ability, we must give all four tests. Gates gives the following example of the scores of four pupils on the four types of tests:

	Tests			
Pupil	*A*	*B*	*C*	*D*
1	100	100	100	100
2	68	60	30	70
3	95	90	40	100
4	75	65	90	100

The figures are the percentages of accuracy. Gates makes the following comments about these four pupils.

"Pupil 1 is a thoroughly accurate reader. In all four tests he made not a single error in comprehension. Pupil 2, of about the same age, grade, and intelligence, is exceedingly inaccurate in his grasp of the material. In Tests A, B, and D, he errs on three or four exercises in every ten and in Test C (Directions) he misses seven out of ten. Whatever else may be discovered about these two pupils, the fact will remain that Pupil 1 is accurate in all types of reading and therefore not in need of cautions and drills in accuracy; whereas Pupil 2 is inaccurate in all types and is in need of help in this respect. Pupil 3 is fairly accurate in reading save in that type which requires very precise and well-integrated understanding, as in Test C. Pupil 4 is very accurate in getting details, either rigidly related ones as in Test C or more independent ones as in Test D, but he is inaccurate in the more general interpretation of a paragraph as a whole, as required in Tests A and B. Pupils 3 and 4 illustrate the value of using the team of tests. Had only Test D been used, both would have been diagnosed as "perfectly accurate readers," whereas, in fact, Pupil 3 is very inaccurate in exacting reading (Test C) and Pupil 4 is very inaccurate in reading to predict something beyond the substance given (Test B) and in reading for the general impression (Test A). On the basis of one test, it is unwise to make a statement concerning a pupil's general accuracy. The team of four gives much greater insight."

5. *Motivation.* — Educational test scores may act as a splendid method for motivating children's work. Pupils may be allowed to keep a record of repeated standard tests in any given subject, so that they may measure their own progress. If they put these on a chart and draw a graph, they are virtually constructing their own learning curves. As we have seen in our general discussion of the laws of learning, work without knowledge of results is never as good as work with knowledge of results. Success is satisfying and a record of progress will therefore help to strengthen the desirable bonds and eliminate the undesirable ones. Nothing succeeds like success. One success spurs us on to a second. This kind of motivation is on general principles to be preferred to the motivation caused by rivalry with others. Although the latter has its merits, yet rivalry with oneself, trying to beat one's own record, is more desirable.

Scales such as the Handwriting Scale are often displayed in the classrooms and children are encouraged to rate their own handwriting and strive to reach a higher level. The average score of the whole class on any test may be used to foster group loyalty and each individual be made to feel his share in helping or hindering the general progress of the group.

So far in educational work the use of tests for motivation has not received much attention. Undoubtedly their value in this direction is great and will be increasingly recognized. Each teacher must use good judgment in such use. To make the child keep too many records at one and the same time would lead to confusion. To continue a record for too long a period may lead to lack of interest. To strive to reach an unattainable goal would not be good motivation. The use of tests for purposes of motivation will only be justified so long as they really continue to function as such.

6. *Research.* — To attempt any record of the uses of tests in educational research would resolve itself into the impossible task of recording the major portion of educational research during the past fifteen or twenty years. The old-fashioned dependence upon the opinion of authority in matters educational is rapidly giving way to the exact methods of quantitative measurement.

The relative merits of various methods of teaching must be investigated by means of tests. The problem of the amount of transfer of training from any one subject to any other can only be satisfactorily handled with the help of tests. In the numerous experiments requiring experimental and control groups, tests are necessary to equate such groups. And so on through the long and varied list of modern educational research we note everywhere the value of educational tests.

Summary

1. The interpretation of educational tests must always be in terms of the intelligence of the pupil.
2. The A.Q. technique is a method whereby the I.Q. and the E.Q. of the pupil are used to determine his general efficiency.
3. Another such method is the use of T scores on educational and intelligence tests in order to arrive at F or efficiency scores.
4. All such techniques assume a similar standardization group from which the different norms have been derived.
5. Six general uses of educational tests have been discussed.
6. General surveys of school systems employ the survey type of educational test to obtain a general measure of the educational product of the system.
7. Efficiency of instruction is best measured by a combination of intelligence and educational tests, in order to evaluate the work of the teacher in terms of the intelligence of his pupils.

8. Educational tests can be used to divide a grade into more homogeneous sections, to help in grade placement of new pupils, to help in deciding on doubtful cases for promotion.
9. Individual difficulties and peculiarities may be discovered by appropriate educational tests. This is called the diagnostic use of the tests.
10. Records of tests may be used to motivate a pupil's work. Rivalry with self is a better motive than rivalry with others. All pupils, good and poor alike, can make progress.
11. For all kinds of accurate educational research the educational test is indispensable.

Review

True-False Statements

As a review mark the following statements true or false:

1. Achievement tests should be substituted for the usual school examination.
2. A high A.Q. means that the child is working too hard and he should be forced to rest.
3. A teacher should be held accountable for the low A.Q. of any child in her class.
4. For survey purposes a relatively short test may be sufficient.
5. Achievement tests are of great help in the proper placing of children in high school classes.
6. The Thorndike-McCall Reading Test would be a good test for diagnosing the specific reading difficulties of any child.
7. An A.Q. above 100 means that a child is working beyond his intellectual capacity.
8. Use of the A.Q. technique is best restricted to tests standardized upon the same population.
9. The best tests for diagnostic purposes are those which are detailed and specific.

Advised Readings

McCall, W. A.: *How to Measure in Education,* New York, 1922, Chapters I to VI.

Strayer, G. D. (Director): *Report of the Survey of the Schools of Port Arthur, Texas.* Bureau of Publications, Teachers College, Columbia University, 1926, Chapters V and VI.

References

1. Gates, A. I.: *The Improvement of Reading,* New York, 1927, Chapter VIII.
2. Strayer, G. D. (Director): *Report of the Survey of the Schools of Tampa, Florida.* Bureau of Publications, Teachers College, Columbia University, 1926, Chapter VIII.

CHAPTER XVI

NEW-TYPE EXAMINATIONS

Classroom Examinations. — One of the functions of a teacher is to score or grade the work of the pupils in the particular subject they are being taught. The teacher during the course of the year attempts to measure the specific modifications he has been trying to bring about in his pupils. From the standpoint of educational psychology the teacher's examinations are measures of the learning progress of the pupils. These classroom examinations differ from standard educational tests, because they are made up by the teacher himself with reference to the specific things he has been teaching, because there are no general standards or norms and thus no outside criterion of the worth or value of an answer or paper. Nevertheless, these classroom examinations are important and, indeed, necessary for our present system of promotions and credits. They are also useful to the teacher himself in order to find out how effectively his teaching is being grasped by the pupils. It is obvious that the more reliable such examinations can be made, the more efficiently will they serve all these purposes.

The Essay Type of Examination. — Until recent times the most commonly used type of examination was the essay type. This type of examination expected the pupils to write an answer in essay form to some question beginning, "Discuss," "Describe," "Criticize," "Tell about," "Compare," "Contrast," "Evaluate," or sometimes more directly beginning with "What," "Why," "How," "Who," or "Where." An examination may consist of one such question or a great many, depending upon the length of the answers expected. In the scoring of such papers the examiner generally gives his judgment upon the question as a whole, assigning a mark of some sort to denote this judgment. If there are several questions the total of these marks represents the total mark for the paper. The standards for scoring are purely subjective and rarely does a teacher decide beforehand just what are the specific requirements for a perfect answer. Usually the examiner depends upon the total impression made upon him by an answer and assigns a mark accordingly. In grading papers for a large class, say a college class of 100 or more, the difficulty of keeping any standard constant in one's mind is very great, and the monotony of the whole procedure is well known. Hence this essay type of examination is difficult to score, it is monotonous and wearisome to grade and it represents a burden which should not be borne by the teacher.

The Essay Type is Unreliable. — In addition to what we have said in the last paragraph about the undesirability of the essay type of examination, there is additional criticism that it is unreliable. The scoring of such examination papers is unfair to the student. There cannot be much significance in a mark if competent teachers of English assign marks all the way from 60 to 98 to the same paper, as has been found to be the case. Again in scoring a history paper, 115 competent teachers assigned marks all the way from 70 to 100, and to another paper gave marks all the way from 45 to 90. In scoring one answer to a question in geography 91 teachers assigned marks from 2 to 20, the maximum score allowed being 20. Such reports as these created dissatisfaction with the essay type of examination and, thus, long before the actual coming of the new-type examination, the need for something better was felt.

The Rise of the New-Type Examination. — The greatest stimulus to the new-type examination was the rapid spread of group intelligence and standard educational tests. Undoubtedly here and there teachers had made occasional use of similar devices in an unscientific tentative fashion.* But the methods now used were first developed in connection with group intelligence and educational tests.

The Types of Objective Examinations. — There are many definite types of objective questions and a brief survey of the more common types will be profitable.

The True-False Type. — Here a statement is presented and the pupil must indicate in some manner whether the statement is true or false.

1. Indicate by encircling T or F:
   T F The normal pulse rate is about 70.
2. Underline true or false:
   Any chord passing through the center of a circle is called a diameter. True False
3. Write "true" or "false" or T or F before the statement :
   ————. The attack on Fort Sumter helped the Northern people, for it united them as they had not been united before.

The statements set before the student may range all the way from mere information to such as require judgment and thought. Statements may be constructed which require considerable critical ability on the part of the student, if he is to answer them correctly. The preparation of difficult, thought-provoking statements is by no means easy. The examiner must guard carefully against ambiguity. The scoring is either the number right or else the number right minus the number wrong. The latter penalizes random guessing. Students should be told whether or not to guess those questions they are uncertain about. Perhaps the best results are obtained by telling pupils not to guess.

* The writer remembers that one of his high school teachers made use of the one word answer type of question in testing for factual information in history, geography, and literature before educational or intelligence tests had been heard of.

The Multiple Response Type.— Here a statement is made with two or more answers or completions, and the pupil is called upon to indicate the best or right answer. Usually at least three choices are offered and the best tests offer four or five choices. A good examination should contain at least four choices so that the influence of chance correct answers by guessing will be very slight.

Underline the correct or best answer:

1. The most important product of Chile is: (1) gold; (2) nitrates; (3) cattle; (4) wheat.
2. The architectural style of the great cathedral at Amiens is: (1) Corinthian; (2) Gothic; (3) Ionic; (4) Doric; (5) Roman.
3. The Romantic movement was interested in: (1) the spread of religion; (2) the life and ideas of the middle ages; (3) the problems of science; (4) the maintaining of the existing order of things; (5) a return to the Roman Catholic church.

The Completion Type. — Here one or more words in a sentence are omitted and the pupil is asked to fill in words so as to make the best or truest statement. Examples are as follows:

1. ———— is the capital of France.
2. Robert E. Lee, of ————, was ———— of the ———— army.
3. Boyle's Law states that when a ———— is subjected to compression and kept at a constant temperature the ———— is ———— proportional to the pressure.

The possibilities of the completion type of question are very great. The questions may range from the very simplest to extremely hard ones. The objection to this form of examination is the difficulty of scoring. The scoring is less objective than that of many of the other types. Many degrees of goodness may be represented in the answers given, and the scorer must call upon his subjective judgment in marking these answers.

The Matching Type. — Here two lists are presented to the pupils and the problem is to match an item of the first list with the appropriate item of the second. For some types of information this is a very good exercise.

Place the number of the author opposite the appropriate book:

Authors	*Books*
1. Burns	—— Eve of St. Agnes
2. Tennyson	—— Mill on the Floss
3. Byron	—— Oliver Twist
4. Goldsmith	—— Idylls of the King
5. Keats	—— Deserted Village
6. Thackeray	—— Cotter's Saturday Night
7. George Eliot	—— Kenilworth
8. Scott	—— Prisoner of Chillon
9. Dickens	—— Lays of Ancient Rome
10. Macaulay	—— Vanity Fair

Such lists may be made more difficult by not having the same number of items in each column. It is best to make the first list a little longer than the second.

The Analogy Type. — Here the student is presented with two words in a certain relationship and he must then find the appropriate word that will fit a third word in the same relationship, or in other words he must complete the analogy. Examples are:

1. Automobile is to carriage as motorcycle is to (1) horse, (2) walking, (3) buggy, (4) bicycle, (5) train.
2. Circle is to square as sphere is to (1) circumference, (2) cube, (3) round, (4) corners, (5) ball.
3. Anger is to fighting as fear is to (1) rage, (2) attack, (3) elation, (4) flight, (5) pain.

Preparation and Use. — The samples above are merely the most common types of the new objective examination. There are other types and teachers may adapt any type to suit their convenience. It will be seen further that the construction of proper objective examinations is much more difficult than the construction of the old-fashioned essay-type examination. We cannot do better than quote here the suggestions of Paterson (1) with reference to the preparation and use of objective questions:

"1. Questions covering every phase of the course should be utilized to insure wide sampling of pupil knowledge.

"2. An excess number of questions should be prepared to allow ample opportunity for the selection of the best questions for the examination proper.

"3. Ambiguous questions both with respect to meaning and possible answer should be rejected.

"4. The apparent difficulty of a question should not be the basis for either accepting or rejecting a proposed question.

"5. Acceptable questions should include an equal number of easy, hard, and moderately difficult questions.

"6. The first half dozen or so questions should be so easy that practically all can answer them, thus serving as a 'shock absorber.'

"7. Each acceptable question should be an independent unit in the examination.

"8. Each acceptable question should be short.

"9. The examination should include a very large number of questions.

"10. Each form or type of question should be segregated, the examination consisting of as many parts as there are types of questions.

"11. Within each part of the examination the questions should be arranged according to topical sequence in the course.

"12. The examination itself should be preceded by suitable general directions.

"13. Specific directions should be given for each segregated group of questions.

"14. There should be a random arrangement of true-false questions, with approximately an equal number of true and false statements.

"15. The correct answers among the alternative answers in the single-choice and in the plural-choice questions should be placed according to chance.

"16. A uniform method of marking the papers, together with the use of a colored pencil in scoring, should be used.

"17. Scoring formulæ should not be used except possibly for the true-false type of questions, when a right-minus-wrong scoring formula may be used.

"18. 'Weighting' of questions according to difficulty or importance is rendered unnecessary in new-type examinations.

"19. Total scores should be computed for the examination papers, distributed on a graph or table, and then a key for converting total scores into letter grades derived.

"20. The examination should be mimeographed or printed, and both used and unused copies should be kept under lock and key in order to avoid the possibility of coaching.

"21. The prevention of coaching should also be accomplished by using duplicate forms for classes in the same subject taking the examination at different hours or on different days and by changing the examination questions from semester to semester.

"22. A large file of questions should be developed for each course, so that a reservoir of from 1500 to 2000 objective questions would be available from which examinations in endless variety could be quickly assembled and used as occasion demands. The ideal plan is to determine the diagnostic significance of each question, thus developing a large list of valid questions to be used in the preparation of examinations."

Advantages of New-Type Examination. — The great advantage of this type of examination is the reliability of the scores, because of the objective system of marking. Reliabilities can be calculated in the usual way by the coefficient of correlation. Furthermore, the marks of such examinations lend themselves more readily than do the marks of the old-type examination to adequate statistical treatment. We may, if we wish, compute T scores for our class and compare the progress of the class or the pupil from time to time with a great deal of reliability. The distributions of marks will tell us whether our examinations are adequately testing the class or whether they are too easy or too hard. In short, the marks of adequate new-type examinations give us a reliable measure of the learning progress of our pupils.

There is no need to discard the old essay-type of examination entirely. It can still be used with profit, but it should be used sparingly, and the teacher should be aware of the difficulties in marking which it presents.

Summary

1. Class-room examinations are the teacher's measure of the particular modifications that he is trying to bring about.
2. The old essay-type examination depends too much upon the subjective impression created in the examiner.
3. This type of examination has been shown to be very unreliable.
4. The rise of the objective examination is due to the development of techniques in connection with group intelligence and educational tests.
5. Useful types of the objective examination are True-False, Multiple-Response, Completion, Matching and Analogy.
6. In the preparation and use of these examinations many details must be watched.
7. The great advantage of the new-type examination is its reliability as a scientific measure of the progress of the pupil.

Review

True-False Statements

As a review mark the following statements true or false:

1. Teachers should not be allowed to use the essay-type of examination any more.
2. The best type of objective examination is the true-false type.
3. A new-type examination should contain easy, medium and difficult questions.
4. The marks given by different teachers to the same essay-type of examination paper have been found to vary enormously.
5. One advantage of the new-type examination is its greater reliability.

Advised Readings

Paterson, D. G.: *Preparation and Use of New Type Examinations*, Yonkers, 1925.

Russell, C.: *Classroom Tests*, Boston, 1926.

References

1. Paterson, D. G.: *Preparation and Use of New Type Examinations*, Yonkers, 1925, 64–66.

CHAPTER XVII

CONCLUSION

We have now come to the end of this brief and elementary survey of educational psychology. Starting with the thesis that the school child or rather the learning individual is the center of interest for our study, we found it necessary to make an artificial division into the original tendencies of the individual and the modifications made by the school. Any such division is merely a matter of convenience. It helps us to bring order into our thinking. We never find original tendencies and modifications separated or clearly differentiated in any individual. Modifications begin at birth and are continually taking place. Yet at the same time every modification is influenced by the original tendencies of the child. There is, thus, a continual inter-action, which makes any child at any given moment a very complex and difficult subject to sum up. Hence the attempt on the part of psychologists to attack this complex situation from two points of view; from the point of view of the original tendencies possessed by the child and from the point of view of the modifications built upon these original tendencies.

The difficulty of keeping to this division has already been apparent to the student in reading this book. The conditioning of instinctive tendencies might well have been taken up in Part II as an aspect of modification, because it is a change or modification of such tendencies. We have kept this discussion in Part I to show that there are original tendencies very early in life and that they are very modifiable. Again, much of the chapter on the measurement of non-intellectual traits might easily have been placed in Part II, as showing the measurement of non-intellectual traits conditioned by the environment. There can be no clear-cut division. It is a matter of convenience as to where we make the division.

All living, all experience means modification of the experiencing organism. Within certain limits human beings are very modifiable. Each human being again possesses his own amount of modifiability. Some take on changes very readily. Others are very refractory to change. Helping to cause these differences in modifiability are the differences in original tendencies possessed by different individuals. Hence the measurement problem for the psychologist is two-fold. On the one hand he attempts from the measurement of one set of reactions to infer the original tendencies possessed by the individual. And on the other hand he attempts by the measurement of other reactions to infer the amount of modification that has taken place. There is nothing inherently different in the reactions themselves. What inferences as to modifications or original tendencies we may draw from the test-reactions will depend upon the group being tested. Thus we frequently see similar test material used in intelligence and in educational tests. Whether it is good for the measurement of original tendencies or of modification will depend upon the children being compared. For all educational measurement is a comparison between children. Our norms or standards are always derived from what some group of children has done.

All life, all experience, we have said, leads to modification. In civilized society, the school is the most important agency set up by man to make certain definite modifications. The school works intensively upon the raw material of human nature for many years, in order to achieve certain modifications. At times the cynic feels inclined to say that the modifications achieved by the school are mighty few, and so he unconsciously pays a tribute to the power of original nature. We also pay homage to the force of these original tendencies when we say, "human nature cannot be changed," meaning thereby that all modifications made are relatively slight and superficial. On the other hand no one would dream of abandoning the school in civilized society, and the optimists among us look to it very largely to change and reform society. If the school is not doing all it should, their answer is more schooling.

These conflicting modes of thought place before the psychologist a two-fold task — the measurement of original tendencies and the measurement of their modification. The school has always been aware of this problem. From the very beginnings of organized education, we note the tendency to examine, in order to find out what modifications the school has made. And in the educational world of to-day, the school examination plays a large and important part. The other aspect of the situation — the measurement of original tendencies — has never entered very clearly into the minds of educators. Although conscious of the differing abilities of children, they have not taken them into consideration very seriously until the present time. The tests of the educational psychologist are attempting to meet this need.

The problem of the modifiable individual is not only a problem of testing, it is also a problem of investigating how modifications actually take place. And hence a large part of our subject deals with the measurement of modifications as they are taking place under varied conditions. From such experimentation we gain some idea as to how children learn and as to the most favorable methods of learning. Only the general results are treated in a text such as this. The specific application of these results must be worked out for every department of study. This is being done in some cases, and so we speak of the psychology of arithmetic, the psychology of reading and so forth. As a matter of fact, research in educational psychology is so recent, that merely the broad outlines have been established at the present time. The teacher must take these general rules and suggestions and apply them specifically to the problems of the classroom as he meets them in his daily work.

GENERAL BIBLIOGRAPHY

Book, W. F.: *The Intelligence of High School Seniors*, New York, 1922.

Buckingham, B. R.: *Research for Teachers*, New York, 1926.

Carr, H. A.: *Psychology*, New York, 1925.

Colvin, S. S.: *The Learning Process*, New York, 1914.

Freeman, F. N.: *Mental Tests*, Boston, 1926.

Galton, F.: *Hereditary Genius*, New York, 1880.

Gates, A. I.: *Psychology for Students of Education*, New York, 1923.

Gault, R. H.: *Social Psychology*, New York, 1923.

Gesell, A.: *The Mental Growth of the Pre-School Child*, New York, 1925.

Hartshorne, H., and May, M. A.: *Studies in the Nature of Character*, I. Studies in Deceit, New York, 1928.

Hollingworth, L. S.: *The Psychology of Subnormal Children*, New York, 1920.

Hollingworth, L. S.: *Gifted Children*, New York, 1926.

Jordan, A. M.: *Educational Psychology*, New York, 1928.

Koehler, W.: *The Mentality of Apes*, New York, 1925.

McCall, W. A.: *How to Measure in Education*, New York, 1922.

McDougall, W.: *An Introduction to Social Psychology*, London, 1922.

Meumann, E.: *The Psychology of Learning*, New York, 1913.

Morgan, J. J. B.: *The Psychology of the Unadjusted School Child*, New York, 1924.

Norsworthy, N., and Whitley, M. T.: *The Psychology of Childhood*, New York, 1918.

Perrin, F. A. C., and Klein, D. B.: *Psychology*, New York, 1926.

Peterson, J.: *Early Conceptions and Tests of Intelligence*, Yonkers, 1925.

Pillsbury, W. B.: *Education as the Psychologist Sees It*, New York, 1925.

PINTNER, R.: *Intelligence Testing*, New York, 1923.

POFFENBERGER, A. T.: *Applied Psychology*, New York, 1927.

PYLE, W. H.: *The Psychology of Learning*, Baltimore, 1928.

ROBINSON, E. S.: *Practical Psychology*, New York, 1926.

RUCH, G. M., and STODDARD, G. D.: *Tests and Measurements in High School Instruction*, Yonkers, 1927.

RUSSELL, C. R.: *Classroom Tests*, Boston, 1926.

SANDIFORD, P.: *Educational Psychology*, New York, 1928.

SPEARMAN, C.: *The Abilities of Man*, New York, 1927.

STARCH, D.: *Educational Psychology*, New York, 1927.

SYMONDS, P. M.: *Measurement in Secondary Education*, New York, 1927.

TERMAN, L. M.: *The Intelligence of School Children*, Boston, 1919.

TERMAN, L. M., et al.: *Genetic Studies of Genius*, Volume I, Stanford University Press, 1925.

THOMSON, G. H.: *Instinct, Intelligence and Character*, New York, 1925.

THORNDIKE, E. L.: *Educational Psychology*, Three Volumes, New York, Teachers College, Columbia University, 1913.

THORNDIKE, E. L.: *The Psychology of Arithmetic*, New York, 1922.

THORNDIKE, E. L., et al.: *The Psychology of Algebra*, New York, 1923.

THORNDIKE, E. L., et al.: *The Measurement of Intelligence*, New York, Teachers College, Columbia University, 1927.

THORNDIKE, E. L., et al.: *Adult Learning*, New York, 1928.

TRABUE, M. R.: *Measuring Results in Education*, New York, 1924.

VAN WAGENEN, M. J.: *Educational Diagnosis*, New York, 1926.

WATSON, G. B.: *Experimentation and Measurement in Religious Education*, New York, Association Press, 1927.

WATSON, J. B.: *Behavior*, New York, 1914.

WATSON, J. B.: *Psychology from the Standpoint of a Behaviorist*, Philadelphia, 1924.

WOODROW, H.: *Brightness and Dullness in Children*, Philadelphia, 1919.

WOODWORTH, R. S.: *Psychology*, New York, 1921.

APPENDIX I

THE ASSOCIATION TEST

Do not read or glance at what follows until you have taken the test described in Chapter III, page 56. If you read this before taking the test, you may as well not take the test, as your results will be perfectly useless either to yourself or to anyone else.

If you have carried out the instructions as laid down in the text, you are now ready to score your paper. Below is the list of 18 stimulus words followed by two response words. If you have responded to the stimulus word by either of the two response words, put a check mark on your papers. Check only the common responses given below:

Stimulus Word	*Common Responses*	
1. Color	Red 61	Blue 26
2. Furniture	Chair 80	Table 8
3. Flower	Rose 61	Violet 10
4. Letter of Alphabet	A 76	B 8
5. Metal	Iron 46	Gold 29
6. Historic Personage	Washington 50	Napoleon 14
7. Part of Speech	Noun 46	Verb 36
8. Geometrical Figure	Triangle 43	Square 23
9. Verb	Run 27	Go 24
10. Tool	Hammer 46	Saw 14
11. Article of Food	Bread 52	Meat 7
12. Part of Body	Arm 35	Head 23
13. Day of Week	Monday 46	Sunday 20
14. Room in House	Parlor 28	Dining 18
15. Animal	Dog 37	Horse 27
16. Book	Psychology 12	Bible 6
17. Girl's Name	Mary 20	Helen 9
18. Country	America 27	U. S. 20

The numbers following the response words show the percentage of students in university classes, that responded with the words in question. These percentages are simply put there as a matter of information.

Now count up the number of common responses which you obtained in the experiment. We shall call that your score. Let us now compare your score with scores obtained by other students who have taken this test. The writer has given this test to many groups of students studying educational psychology. Here are the results. Q_1 means the lower quartile and shows the score below which 25 per cent of the class fell; m is the median score — the score made by the middle person in the class, half of the scores lie above and the other half below this median score; Q_3 is the upper quartile and 25 per cent of the class lie above this score; the range shows the lowest and highest scores obtained by any member of the class; n means the number of students in the class who took the test.

Class	Q_1	*Median*	Q_3	*Range*	n
1.	10	12	14	4–17	119
2.	9	12	13	4–17	59
3.	10	12	14	5–17	262
4.	11	12	13	4–18	93
5.	11	13	14	6–17	130
6.	10	12	13	4–17	131
7.	9	10	12	6–16	41
8.	10	12	13	6–18	37

In general the middle 50 per cent of students in a class are likely to have scores between 9 and 14. If your score lies between 9 and 14, then you have reacted in the most normal or ordinary way for students of educational psychology in general to react. If your score lies above 14, anywhere from 15 to 18, then you have responded with more common responses than is usual. You belong among a group of 25 per cent of students who give these higher scores. As your score tends to reach 18, you become more and more unlike the mass of students in your associations; you have too many common responses. If your score lies below 9, anywhere from 4 to 9, then you belong to the 25 per cent of students who give few common responses. If your score should be below 4, then you have fewer common responses than any of the 872 students I have so far tested.

This is just a little experiment to give you an idea of how association tests work. Do not be alarmed if you have given too many or too few common reactions. I am comparing your scores with results obtained by normal university students. There is here no question of investigating abnormality or insanity. Just what the significance of the number of common responses is, it is rather hard to say. Perhaps the more sociable we are, the more we think alike. The less sociable, the less likely are we to think like our fellows and the more likely are we to develop thoughts and associations somewhat different from our fellows. One of my students tried to follow up this idea and obtained a rating of "sociability" from the students themselves and from other students. When she correlated these ratings with the number of common associations on this test she found a positive correlation of .45 with the self ratings and of .33 with the ratings by others.

Just why we react with a certain word, it is frequently very difficult to say. If you have a class discussion on the results of this test, you will be surprised at the different kinds of reasons that the students will give for the specific reactions they have made. Some of the reactions can be explained by frequency, that is, they are associations between words or ideas that are very frequently used, such as "furniture—chair," or "geometrical figure—triangle." Because they are the first of a series probably accounts for "letter—A," and "day of week—Monday." Note in the latter connection that Monday is generally the most frequent response regardless of the day on which the test is given. "Book—psychology" indicates the general nature of the environment under which the test was taken, but remember that the day of the week did not influence the response. It is interesting to note further that Napoleon is thought of as an historic personage more frequently than Lincoln, that many girls respond to "girl's name" by "Mary" or "Helen" and very rarely do they give their own name. Every now and again some student will give one or more responses of an uncommon type that are directly connected with some striking event that has just occurred to him. See if you can find samples of such reactions in the class, but be critical about accepting all of the explanations offered. I have known students to say that they responded to "historic personage" by "Napoleon," because they had just been studying or reading about Napoleon. I doubt such an explanation, in view of the fact that in general about 14 per cent of students give such a reply.

This test can be given to children in the grades. You can try it with classes of school children, if you are a teacher. It is surprising how early in life these common associations seem to be formed. Here are the commonest responses given by 236 older children, age thirteen and up, and by 119 younger children below age thirteen:

Stimulus	*Older Children*	*Younger Children*
1. Color	Red	Red
2. Furniture	Chair	Chair
3. Flower	Rose	Rose
4. Letter of Alphabet	A	A
5. Metal	Gold	Gold
6. Historic Personage	Washington	Washington
7. Part of Speech	Noun	Noun
8. Geometrical Figure	Triangle	Square
9. Verb	Run	Run
10. Tool	Hammer	Hammer
11. Food	Bread	Bread

Stimulus	*Older Children*	*Younger Children*
12. Part of Body	Arm	Head
13. Day of Week	Monday	Monday
14. Room in a House	Kitchen	Bedroom
15. Animal	Horse	Horse
16. Book	History	History
17. Girl's Name	Mary	Mary
18. Country	America	A Town

APPENDIX II

JUDGMENT OF INTELLIGENCE FROM PHOTOGRAPHS

The ranking according to the intelligence tests of the twelve children is given below. The children were tested by the Yerkes-Bridges Point Scale and the intelligence rating is the C.M.A. or Coefficient of Mental Ability, which is not identical to the I.Q. or Intelligence Quotient. C.A. means Chronological Age and M.A. means Mental Age.

Rank	*Child*	*C.A.*	*M.A.*	*C.M.A.*	*Description*
1	C	4	5.7	1.93	Very bright
2	D	5	7.2	1.64	Very bright
3	G	10	12.	1.21	Bright
4	H	10	10.8	1.09	Normal
5	K	7	6.8	.97	Normal
6	L	16	12.5	.90	Normal
7	A	10	9.	.88	Slow
8	B	12	9.7	.81	Slow
9	E	12	10.	.80	Slow
10	F	15	9.8	.72	Very dull
11	I	12	8.6	.72	Very dull
12	J	14	7.6	.47	Feebleminded

The instructor should work out a rank correlation with the class. He should work one example on the blackboard and have the class follow step by step as explained in the text. Here is a sample of how the work should appear on the blackboard and on each student's sheet.

Child	*My Rank*	*Test Rank*	*Difference*	D^2
A	5	7	2	4
B	9	8	1	1
C	3	1	2	4
D	4	2	2	4
E	10	9	1	1
F	6	10	4	16
G	2	3	1	1
H	8	4	4	16
I	1	11	10	100
J	11	12	1	1
K	7	5	2	4
L	12	6	6	36
				188
				6
				1128

$$\rho = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$
$$= 1 - \frac{1128}{1716}$$
$$= 1 - .66$$
$$= + .34$$

The following coefficients of correlation were reported by the writer for different groups of individuals who have taken this test.

Distribution of Coefficients of Correlation

	Physicians	*Psychologists*	*Miscellaneous*	*Students*	*Teachers*
	28	45	52	29	37
	25	41	39	27	34
	21	40	31	24	33
	19	38	14	15	27
	04	34	09	06	27
	−08	33	01	03	24
	−10	31	01	−06	15
	−14	27	01	−07	13
	−21	15	−19	−08	10
		13	−32	−24	08
		−01	−51	−30	08
		−02			06
		−09			−05
		−63			−07
					−26
					−28
					−29
n..........	9	15	11	11	17
Av. r.......	+.05	+.18	+.05	+.03	+.09
Av. rank r..	+.11	+.31	+.06	+.03	+.17

Median r for all cases = + .10.

A full description of the original experiment will be found in the following: Pintner, R., "Intelligence as Estimated from Photographs," *Psych. Rev.*, Vol. 25, No. 4, July, 1918, pp. 286–296.

Below is given another distribution of correlation coefficients obtained from a large class of graduate students in education who made their judgments from a screen projection of the same pictures.

Distribution of Correlation Coefficients

Coefficients	*No. of Cases*
Plus 63 to plus 87	9
Plus 38 to plus 62	9
Plus 13 to plus 37	28
Minus 12 to plus 12	30
Minus 37 to minus 13	10
Minus 62 to minus 38	4
Total	90

Median Coefficient = + .14

INDEX OF SUBJECTS

INDEX OF NAMES