ia

EDUCATIONAL PSYCHOLOGY

VOLUME III

MENTAL WORK AND FATIGUE
AND
INDIVIDUAL DIFFERENCES AND THEIR CAUSES

EDUCATIONAL PSYCHOLOGY
VOLUME III

MENTAL WORK AND FATIGUE
AND
INDIVIDUAL DIFFERENCES AND THEIR CAUSES

BY
EDWARD L. THORNDIKE
PROFESSOR OF EDUCATIONAL PSYCHOLOGY IN TEACHERS COLLEGE,
COLUMBIA UNIVERSITY

PUBLISHED BY
Teachers College, Columbia University
NEW YORK
1926

MASON PRINTING CORPORATION
SYRACUSE NEW YORK

PREFACE

This third volume presents the results of psychological studies of mental work and fatigue (in Part I), and of individual differences and their causes (in Part II). Part II is a revision of a book, *Educational Psychology*, which appeared in 1903 and, in revised form, in 1910. Part I is entirely new. The same procedure—of introducing topics by means of reports of typical investigations, presenting instructive evidence as well as conclusions, and discussing the important principles of quantitative treatment in each case—adopted in the earlier volumes is maintained.

Teachers College, Columbia University,
November, 1913.

CONTENTS

PART I

PART I

Mental Work and Fatigue

CHAPTER I

MENTAL WORK AND FATIGUE: DEFINITIONS AND PROBLEMS.

The topic of this and the five following chapters is the temporary deterioration of mental functions due to exercise without rest—its amount, rate and changes in rate, the factors constituting it, the conditions by which it is influenced, and the effect of such deterioration in one function upon the efficiency of others.

It is, however, inadvisable to present only the actual facts concerning these divisions of the topic, or to present the facts in exactly the order indicated. Though these divisions and this order are logically the most suitable, a strict limitation of the exposition to them and a strict maintenance of this order would hinder the student's insight into and sympathy with the investigations of mental fatigue which have actually been made.

These investigations were made, for the most part, without any definite understanding of what a mental function or activity or ability should be assumed to be. 'The mind' was supposed vaguely to 'do work' and to 'be fatigued' thereby, and various supposed tests of its efficiency, on the whole or in some particular, under various conditions as to work and rest, were made. Many of the investigators took it for granted that its efficiency in certain particulars would perforce be a mere sign and parallel of its efficiency as a whole. Many of them also assumed that the mind must, without frequent and long rests, become fatigued, and planned their investigations to find convenient symptoms of this fatigue.

To make proper connections for the reader with these investigations and their results, it is necessary to accept to some extent and provisionally this older point of view, to get working definitions of 'mental work' and 'mental fatigue,' and to rehearse the results of certain tests before and after a period of mixed mental work (such as a school-session) though these do not bear directly on any of the sub-topics listed. Also there has been so little study of the factors in, and conditions of, the temporary diminution of a mental function exercising with inadequate rests, that these topics cannot be given the space which their importance warrants.

On the other hand, to enable the student to know in each case what fact was really measured and what the measurement really meant concerning any defined function or ability, it is necessary to replace the vagueness of 'mental energy,' 'mental work,' 'mental fatigue' and the like, just as rapidly and fully as possible, by more workable concepts, to bring together the ascertained facts concerning the amount and rate of deterioration of a function from lack of rest, and to do the same for changes in the rate, and for the 'transfer' of the deterioration to other functions.

The compromise made here between fair presentation of studies that have been made and systematic presentation of the facts concerning the temporary effects of insufficient rest upon mental functions results in the following topics and order:

Mental Work *versus* Bodily Work.

Mental Work *versus* Mental Play and Mere Activity.

The Concept of Mental Fatigue.

The Deterioration of a Single Function under Continuous Exercise.

The Curve of Work.

The Curve of Satisfyingness.

The Influence of Continuous Mental Work, Special and General, upon General Ability.

The Symptoms of Mental Fatigue.

Theories of Mental Fatigue.

The Hygiene of Mental Work.

The discussion under each topic will be as straightforward as I can make it without subjecting the results that have been attained to too great reorganization.

MENTAL WORK *versus* BODILY WORK

Mental work is some part of the total work done by an animal. This total may be divided into mental work, muscular work, glandular work, the work of the red blood corpuscles, the work of the white blood corpuscles, the work of various enzymes, and so on. Just as no exact limits have been set between the behavior due to 'the mind' and the behavior due to 'the body,' so no exact limits have been set between the mental and the non-mental work of an animal. Such work as obtaining the answers to problems in arithmetic, physics, law or pathology, all would agree in calling mental. Such work as transforming starch into dextrine, killing off malarial parasites or pumping the blood, all would agree in calling non-mental. But, as we shall see, there are operations, such as playing the piano or sewing, which are troublesome to classify.

There are four important possible criteria for marking off the mental part of the work done by an animal. The first is the production of conscious states. The second is the presence of conscious states, whether produced by the work or not. The third is the absence of explanation by present physiology. The fourth is action by the animal's connection-system.

To define mental work as the work of producing states of consciousness is not useful. For, in the first place, the amount of consciousness produced is relatively insignificant in comparison with what particular sort of consciousness is produced and in what connection. Much intellectual or moral achievement may go with little consciousness, as in directing an army; and much consciousness, with little achievement, as in revery. In the second place the control of certain movements is as important a feature of intellect as the control of thoughts.

Nor is it useful to define mental work as achievement accompanied by consciousness. For the work of swallowing food, holding up a weight, or excreting tears would then often have to be classed as mental; and the work of the telegrapher in translating series of clicks into written words, or of the actor in playing his part, would at times have to be classed as non-mental. The same effective response to the same external situation would be mental work, partly mental work, or non-mental work, on different occasions. Moreover, it is doubtful whether the really essential feature of mental work, the production and prevention of connections, is itself accompanied by consciousness.

To define mental work as that portion of the animal's work which physiology does not account for, is, for the present, reasonable. By tradition physiologists have assumed certain tasks and left others to psychologists, sociologists, economists and historians. The word *mental* in common scientific usage does refer approximately to what they have so left. The physiologist starts at one extreme with the work of supporting weight or absorbing oxygen; the student of human affairs starts at the other extreme with the work of managing enterprises, doing sums, writing books, or memorizing facts. The ground not covered by either one may be allotted to the other. Such a definition is obviously provisional. If there is any gap somewhere in the series of intermediate forms between absorbing oxygen and writing a book, it will be a better basis for the division than the casual division made by the interests of men of science. If there is no such gap, if the forms of an animal's work vary continuously from giving forth CO_2 to giving forth poetry, then mental work should be frankly recognized as differing from other animal work only as one extreme of a continuous series.

The fourth means of distinguishing mental from other animal work has the merits, first, of being clear, and second, of substantially following the best present practice of students of mental work. To define mental work as work done by the animal's connection-system leaves no danger of confusion with the work of the digestive, excretory, action, or other systems. There is then no such obscurity about whether a given piece of work by a given animal is to be called mental or not as exists if the presence of consciousness, accompanying the work or produced by it, is made the criterion. The work of the connecting neurones also does include almost all of the work which expert opinion would unanimously call mental, and exclude almost all of the work which expert opinion would call non-mental. Inference, analysis, memorizing, skill in technique and craftmanship, 'observation' and 'attention' are all functions of the connecting neurones. The work of gland-cells in excreting, muscle-cells in contracting, white blood-corpuscles in killing bacilli and the like, which are clearly excluded, comprise the typical forms of 'bodily' or 'physiological' or 'non-mental' work.

The cases of work sometimes called 'mental' which cannot be attributed to the connecting neurones, and the cases of work done by the connecting neurones which are not always called mental, are instructive.

The term, mental work, is sometimes applied vaguely to an activity which in its entirety involves sensory and muscular action as well as thought, self-control or skill. Adding, for instance, means seeing or hearing and saying or writing, as well as remembering the addition combinations and resisting distractions. Solving geometrical problems gives work to the eyes and fingers as well as to the 'reasoning power.'

Such a total activity, though called mental work rather than a composite of sensory work, intellectual work and muscular work, involves the reception and action systems, as well as the connection system. Diminished efficiency of the activity as a whole may be due to inefficient reception of outside stimuli by the first sensory neurones, or to inefficient conduction of stimuli from the neurones to the muscles, as well as to inefficient conduction along the associative neurones and across their synapses.

When, however, such a total activity is examined more critically, it is found desirable to separate off the work and fatigue of the sense-organ and of the end-plate in the muscle from the work of the connection system and to distinguish sensorial fatigue, intellectual fatigue and muscular fatigue. For the action of a sense-organ or of an end-plate is only partly, and the action of a muscle-fibre is not at all, like that of the connecting neurones. If one term is used for all, there is a risk that something found true of only one will be falsely asserted of the others. Moreover, if the total activity is left without further analysis, one may be misled into futile attempts to control the activity. Consider, for example, the different preventatives that should be applied for fatigue in adding, supposing it to be due (1) wholly to diminished efficiency of the rods and cones in being set in action by light, or (2) wholly to diminished contraction of the eye-muscles in moving the line of regard up and down the columns, or (3) wholly to diminished efficiency in connecting '*thought of 13 and perception of 9*' with '*thought of 22.*'

The work done by the connection system which is not usually called mental work comprises such activities as the regulation of breathing, the heart-beat, digestion, and the like. There is here a real difficulty in classification. The work of the connection system does include the provision of responses which the intellectual and moral nature of an animal never or almost never influences, to situations which that nature never or almost never regards. Between these and the habits of scholarship or citizenship, there is a continuous gradation, so that any dividing line has to be arbitrary. The most useful placing of it to fit present usage could be determined. Or, instead of making any such division, we may find it more useful not to place it at all, but to deliberately widen the usage of 'mental,' making it equal to 'caused by the connection-system.'

As a result of this review, the most useful arrangement of terms seems to be to use 'mental' work for the work done by the connection system, or to replace the term by less troublesome ones. Acts of morals and of skill—that is, of the appropriate connection of movements—would, if the term is thus used, be included. The mere impressibility of the first sensory neurones by light, heat, and so forth, would be left for special treatment under 'Sensory Work and Fatigue;' and, *a fortiori,* the impressibility of the muscle by action at the end plate of the motor neurones, and the power of its fibers to contract, would be left for separate treatment under 'Muscular Work and Fatigue.'*

MENTAL WORK *versus* MENTAL PLAY AND MERE ACTIVITY

In the differentiation of an animal's total work into mental and non-mental, psychologists have assumed that they knew what an animal's total work is. But the exact lines between work and recreation or play, and between work and mere inefficient activity, have not been drawn. For understanding of what has been thought and proved about mental work and fatigue, the vagueness and inconsistencies of common definitions of and distinctions between work, recreation and mere

* It will doubtless be better in the long run to subdivide the work of the human animal still further and to replace the terms *sensory, mental* and *muscular* work by the work *of the accessory apparatus of the sense organs,* the work *of the peripheral end of the first sensory neurone,* the work *of conduction along a neurone,* the work *of conduction across a synapse from neurone to neurone,* the work *of changing the intimacy of synapses,* the work *of conduction from the ending of a neurone to a muscle,* the work *of the muscle fibers*. But for the present it seems best to retain the term mental work, relieved of some of its ambiguities and misleadings.

mental life probably do little harm. But it is worth while to be guarded against certain errors.

There are two main sets of criteria for defining work, the first set referring to the kind of product produced, the second set referring to the way in which it is produced. Certain products of intellect, character and skill—such as essays written, temptations overcome, or pictures painted—are commonly taken as evidence of mental work; and this is in a sense reasonable. However, satisfactory distinctions between mental work, mental play and mere mental life cannot be made on the basis of the result achieved. The same desirable result—for example, a given poem or answer to a problem—may come from activity against resistance, from activity along a line of no resistance, or from a mere casual effervescence of mental life. It may come with effort, with release from effort, or just come. Its production may be laboriously painful, a sheer pleasure, or approximately indifferent. So also, a useless and undesirable result, such as an obsession or superstitious fancy, may come from hard work as well as from playful amusement or mere happening. Computation is usually work, but not to the arithmetical prodigy. Cracking jokes may be painfully laborious to the hard-pressed professional humorist. Wise men play whist; socially ambitious or conscientious women and gamblers often work at it; a regrettably large company of benighted persons sit at its tables leading and following according to the casual pulses of an aimless life.

Certain processes such as attention, controlled association and memorizing are also commonly supposed to testify to work rather than play or the mere flow of life; and this, too, gives a roughly useful division. The essential feature in such processes, which has led to their being used as criteria of work, is the inner attitude of effort which so often characterizes them. Consequently, the distinction of work from play and mere activity by means of the process involved in the activity generally gives way to distinction by differences in the individual's attitude at the time. In proportion as the activity is intrinsically satisfying to a man, he calls it play; in proportion as it is intrinsically annoying, depending upon ulterior, secondary facts for its motive, he calls it work. In proportion as it is an undertaking or a part of an undertaking ordered by man, it is called work or play; in proportion as it is the result of inner or outer impulsions irrespective of man's intent, it is called mere activity, mere life, a mere happening. A large proportion of a man's life is a mixture of the three.

There are thus two independent variables involved in mental work, as the term is ordinarily used. There is the production of a certain product and the overcoming of a certain discomfort.

Every measure of mental work in the popular sense needs to be thought of as a composite of two measures on two scales. We can measure the quantity and quality of certain products,—numbers added, words translated, truths told, letters telegraphed, cities conquered, and the like. We can measure, though very imperfectly, the degree of satisfyingness or intolerability of any given hour. But we have not equated the numbers added in an hour and the intolerability of the hour into a single measure of behavior at an intrinsically unsatisfying task. We may ask, for example, how hard a man worked in a given hour, meaning either (1) "How much computation did he accomplish; how much of a product did he turn out?" or (2) "How much resistance from intrinsic unsatisfyingness did he overcome; how much effort did he put forth?" To either one of these questions taken singly an answer can be got. But under present limitations of knowledge, confusion results if we try to give one single answer to the composite question, "How hard did the man work?"*

We measure one or the other fact; or measure one and

* This is readily seen in such a case as: A, in 100 minutes, added 10,000 numbers, reporting honestly that he had never worked so hard in his life. A, the day previous, added, in 100 minutes, 20,000 numbers, reporting honestly that the work was a mere trifle.

the other and report them separately. The quantity and quality of the product produced is usually the fact to be measured; the resistance overcome in the shape of the intrinsic unsatisfyingness of the process is usually left unmentioned, or very crudely located by common knowledge of the nature of the task. What has commonly been measured as mental work is the amount and quality of certain products whose production is *somewhere on the minus side* of the zero of intrinsic satisfyingness.

In the following discussion the word Work will, except when some other meaning is clearly indicated in the context, be used of behavior undertaken by, not merely happening in, the individual, and such as he would not at the time undertake for its own intrinsic satisfyingness. This is the meaning commonly attached to it in the studies whose results are to be reported.

In the end it will be best to abstract mental achievement entirely from the attitude in which it is achieved, not separating off as work all that production which is on the minus side of the zero of intrinsic satisfyingness, and as play all that which is on the plus side. Mental work would then be defined as the production of intellectual and moral products, whether the producer intended to produce anything or not, and whether he would or would not have avoided the production, in and of itself.

For an understanding of the facts and theories to be dealt with in this volume, all these difficulties of definition will be of little moment. For we shall specify the concrete achievements that are in question in each case; we shall be concerned with work—such as computation, memorizing, reacting to a signal, recalling the opposites of words, or correcting examination papers—the major part of which is mental work by whatever definition; and we shall make whatever distinctions each case requires when that case is before us.

THE CONCEPT OF MENTAL FATIGUE

The term fatigue has been used by psychologists in four different meanings: (1) A temporary decrease in efficiency resulting from work done without adequate rest; (2) A temporary decrease in efficiency resulting from work, play or mental action of any sort without rest; (3) A temporary decrease in efficiency resulting from work, disease, drugs, or whatever cause; and (4) A complex of feelings. Let us reserve the word when used alone for the first and second meanings and always use 'feeling or feelings of fatigue' for the fourth. The third usage is inadvisable; the temporary deteriorations due to the action of disease, insufficient nutrition, drugs, blows, the withdrawal of blood from the brain to the digestive organs, and the like, should be discussed by themselves, except in so far as the disease or lack of nutrition is a secondary consequence of the continued work.

It is a fact of common observation that mental work done without rest often results in a decrease in efficiency—that is, a decreased amount of work per unit of time, or an inferior quality of work, or both. Theories of two sorts have been set forth to account for the fact. The first consider mental work after the analogy of the work of an engine, waterfall or dynamo, supposing a quantity of mental energy to be used up as work is done and reproduced during rest. They may be called *Mechanical Theories* of mental work and fatigue. The second consider mental work as the action of certain situation-response bonds, supposing tendencies unfavorable to their action to be produced as work is done and to die out during rest. They may be called *Biological or Response Theories.*

According to the Mechanical theories, fatigue is *intrinsic* or *direct* and *negative* or *subtracting,* in the sense that an activity in and of itself weakens its own efficiency by being exercised without rest, as a reservoir by discharging water lowers its pressure. According to the Biological theories, fatigue is also *extrinsic* or *indirect* and *positive* or *additive*, in the sense that an activity, by being exercised without rest, produces certain by-products, or releases certain forces, external to it, which check it.

According to the mechanical theories, temporary deterioration is referred vaguely to a loss of some one thing—mental energy. According to the biological or response theory, the deterioration is referred specifically to changes in the strength of certain connections on the one hand, and to changes in the readiness of certain conduction-units on the other.

The mechanical theories are commonly left somewhat vague, and those who have accepted them would probably differ in details, for example, in respect to whether or not each separate amount of mental work done lowers efficiency, as each discharge from a reservoir lowers its pressure. But all agree in supposing that doing mental work uses up some quantity of mental energy so that less is available until rest recreates it.

The biological or response theory has been neglected by the great majority of students of mental fatigue, so that the exact theories to which it will give rise cannot be surely prophesied. But they will agree that mental work without rest decreases efficiency chiefly by increasing resistance, by producing conditions, partly outside of the connections involved in the work itself, which make the results of those connections less satisfying.

I shall return to the discussion of general theories of mental fatigue after the measurements which have been made of the effect of mental work of various sorts upon mental efficiency have been reported. These measurements will enable us to clear up some of the ambiguities, inadequacies and obscurities to which the treatment of this chapter has been subject.

CHAPTER II

The Decrease in Efficiency of a Single Function under Continuous Exercise

The term efficiency is used here to refer to the quantity and quality of the product produced. If the quantity per unit of time remains constant, decrease in efficiency is measured by the decrease in quality; with quality constant, by the decrease in quantity; with both varying, by some composite of the two changes.

The term single function is used in antithesis to 'the mind as a whole,' not to mean a function devoid of compoundness or complexity. I mean by it such functions as adding a column of figures, reacting to a signal by a movement as quickly as possible, the signal and movement being the same throughout, judging which of two weights (all close to 100 grams) is heavier, memorizing the English equivalents of German words, or multiplying a three-place number by a three-place number, nothing being allowed to be written or spoken save the two numbers themselves. Each of these functions comprises different elements, not all of which are at work all the time. Exercise is continuous only in the sense that the subject does his best to make it so.

I shall first describe at some length a sample experiment with a single individual and the information concerning the amount, rate, and change of rate of diminution in efficiency derived from it. I shall then summarize the facts known concerning the amount and average rate of fatigue in single functions.

A SAMPLE EXPERIMENT

As the sample for intensive study, I choose an experiment of Arai ['12], which has the special merit of measuring the effect of continuous exercise of a very difficult intellectual process, as free as possible from sensory or muscular work, at a stage when it was almost free from improvement by practice, over a very long period, and on four different occasions. Miss Arai says:

"The first experiment was made during February and March, 1909, at Teachers College, Columbia University. The purpose of the experiment was to ascertain: (1) the amount, rate and the change of the rate of fatigue in the special mental function exercised, and (2) the amount of fatigue transferred to certain other functions.

"The particular function tested was that exercised in mental multiplication of pairs of numbers like

2645	8324	7954		5438
5784	7384	3528	and	2347
———	———	———		———

"About one thousand different combinations of figures were used. The order of the examples being made by chance, the distribution of difficult and easy examples is random. The subject of the experiment was the writer herself. . . . On February 2nd, the subject made the first test in the following manner. Using an ordinary watch the subject set a time for starting. When the hand of the watch reached the point set, the subject looked at the first example and multiplied mentally but with the original numbers in sight. The answer was written down as soon as it was obtained and the time recorded. Then the subject immediately took up the second example and repeated the same procedure. Thus she worked from 9:30 A. M. to 3:18 P. M. with a rest of forty-eight minutes for luncheon, and obtained the answers of twenty-four examples." ['12, p. 31 f.]

On February 4, twenty-six such examples were done in the same manner; on February 7, twelve; on February 15, thirty; and on February 22, sixty. From the seven hours continuous work of February 22, it was clear that the subject could not, by even this long period of work, be brought to a condition of inability to do the work. The work was consequently made still more difficult, as follows:

"Instead of multiplying with the original figures in sight, the subject relied on memory for the figures and multiplied them mentally with closed eyes. The method was better than the earlier one, for it not only made the task more difficult, but it helped to eliminate sensory fatigue. When the subject forgot the original figures, she looked at them again, but as the time was made longer on this account, the loss of the original figures was counted against her. But this seldom occurred as the subject was careful to commit the numbers to memory." [Arai, '12, p. 35]

Her work, that is, was to look at an example such as $\begin{smallmatrix}4962\\ \underline{7584}\end{smallmatrix}$, cover it, memorize the two numbers, then multiply mentally 4 × 4962, getting 19848, then memorize this, but keeping in mind the 4962 and the 758 to be used later, then to multiply mentally 4962 by 8 getting 39696 and perform mentally the operation of adding $\begin{smallmatrix}19848\\ 396960\end{smallmatrix}$.

Having obtained the 416808,* she could now forget the 19848, but must not have let slip the 4962 and 75 and must remember the 416808. She then multiplied 4962 by 5, and remembered to count the 24810 as 2481000 in adding it to 416808. Having obtained 2897808, she could now forget all but it and the 4962 and 7 and the fact that the 7 counted as 7000 in multiplying. Multiplying mentally and obtaining 34734, she held it in mind and added 34734000 to 2897808, and could then write down the answer 37,631,808, look at her watch, record the time, look at the next example on the sheet, say $\begin{smallmatrix}9653\\ 7267\end{smallmatrix}$ and proceed as above.

If the reader will try this work with the far easier task of multiplying four-place by three-place numbers for even only an hour or two he will appreciate that it is far more

* Other methods of operation are, of course, possible, but this was the one which she used throughout the experiment.

difficult and fatiguing (in the popular sense of requiring much disagreeable effort) than all but a few of life's customary intellectual labors.

After doing 189 examples requiring about thirty-five hours (during the week February 24-March 2) by this new method, the subject reached a point where practice effect was very slight, and secured in the next four days the record to be considered here.

"On March 3, 4, 5 and 6, the subject did the mental multiplication from 11 A. M. to 11 P. M. without any pauses except the two or three seconds between the examples for recording time. But the subject had taken a heavier breakfast than usual at 10 A. M. and a light supper after 11 P. M. Her health was in good condition and she slept soundly at night. The contents of her consciousness during the experiments were very simple, all desires being completely subjected to the one desire to get true fatigue curves." [Arai, '12, p. 37] The results of these experiments are summarized in Table 1 and FIG. 1.*

The base line of FIG. 1 is scaled for the number of examples done, one inch equalling forty examples, each 12 hours of rest being denoted by a quarter-inch vertical line at the appropriate point on the base-line.

Above each tenth of an inch along the base line a horizontal line is drawn whose height in each case denotes the time

* There are discrepancies between the measures given here and those given in Miss Arai's published report. Both were taken from Miss Arai's original written records, and either she or I made two mistakes in copying. As I am unable now to find out which of us is in error, I let my figures stand. This difference is unimportant. Also I have attached a much heavier penalty for inaccuracy than she did.

The allowance in time for errors made is of course arbitrary, but any reasonable allowance would give substantially the same general curve. Since the height of the curve formed when these horizontals are joined measures approximately the time that was or would have been required for successive groups of four examples, done with equal accuracy, fatigue is shown by the extent to which the curve rises with work and falls with rest.

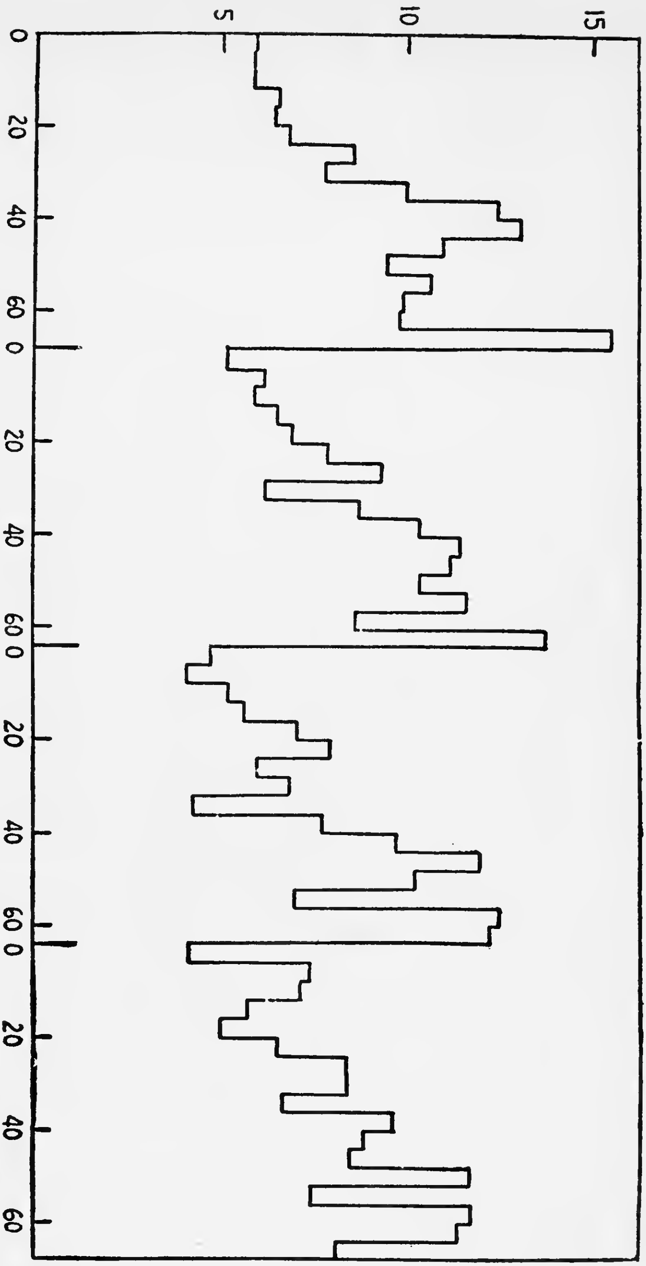


FIG. 1. Curve of Work in Mental Multiplication with Four-Place Numbers. Each inch on the base-line equals forty examples done. The height of the curve represents the time required, with allowance for errors. The rests between the end of one work-period and the beginning of the next are shown by quarter-inch verticals above the zeros on the base-line.

required for the four examples in question, plus an addition for each wrong figure more than two in any answer, and a subtraction for each wrong figure less than two in any answer, of three per cent of the time required for the four examples (that is, twelve per cent of the time required per example).

TABLE I.

FATIGUE IN THE CASE OF MENTAL MULTIPLICATION WITH FOUR-PLACE NUMBERS. After Arai, '12, p. 38 f.

Time required (in minutes), with allowances for errors, for successive sets of four examples multiplied mentally.

Set	Mar. 3	Mar. 4	Mar. 5	Mar. 6	Average
1	23.6	20.7	19.3	16.5	20.0
2	23.3	24.5	16.5	29.6	23.9
3	23.2	23.5	20.9	28.5	23.4
4	26.1	25.9	22.8	23.0	24.6
5	25.8	27.8	28.3	20.2	26.8
6	27.3	31.4	31.7	26.2	29.4
7	34.3	37.3	24.0	33.6	34.0
8	31.3	24.9	27.5	33.8	29.4
9	40.0	35.0	17.1	26.7	30.9
10	49.8	41.5	31.0	38.6	40.0
11	52.2	45.8	39.1	35.6	42.5
12	43.8	44.6	48.1	34.1	44.2
13	37.9	41.8	41.0	47.0	41.4
14	42.5	46.5	27.9	29.8	36.2
15	39.7	31.1	28.3	47.1	36.6
16	39.0	52.0	50.0	45.6	46.7
17*	62.1	44.4	49.1	32.9	47.1
First eight	46.9	45.2	35.8	46.1	43.9
Last eight	101.1	96.4	99.1	78.5	93.8

The Amount of Fatigue

The amount of fatigue is measurable by the increase in time required† as work continues, an allowance being made

* There were only three examples in the 17th set. The score was adjusted to be what it would have been for four examples done at the same speed and accuracy.

† Here and throughout the discussion of Miss Arai's experiment *'time required'* means *'time required for work of equal accuracy.'*

for practice, or by the increase of time required at the end of work over that required *at the beginning of work after full rest.* Fatigue then is shown (1) by an average increase of time for the last eight over the first eight examples on days 1, 2, 3 and 4 of 119%, plus whatever correction should be applied for the practice gain. Using only days 1, 2, and 3 the percentage is 135. The practice gain may be estimated by the change from day to day, especially in the first hour of the work. There was, on account of the large amount of preliminary practice, little or no practice effect observable within these four days, the times required for the first sixteen examples on successive days being 96, 95, 90 and 98 minutes; and for the first eight examples, 47, 45, 36 and 46. Or fatigue is shown (2) by an average increase of the time required for the last eight examples of days 1, 2 and 3 over the first eight examples of days 2, 3 and 4, of 136%. This figure would probably be lowered if a comparison of the last eight examples of day 4 with eight done twelve hours later after rest could be included. There was least decrease in efficiency on the fourth day.

This, it must be remembered, by no means implies that the function was less than half as efficient at the end of the twelve hours of work as at the end of twelve hours of rest. On the contrary, the amount of percentile loss in absolute efficiency was probably very slight. For a person to be able to multiply a number like 9263 by one like 5748 without any visual, written, or spoken aids, even in fifteen, or for that matter in a hundred and fifty, minutes, implies a very high degree of efficiency. That a person can exert himself to the utmost at this very difficult work for ten or twelve hours without rest and still be able to do it, even if at the expense of twice or thrice as many minutes per example as at the beginning, means that the loss in efficiency by any absolute standard has been small. For Shakespeare to have required twice as long to write Hamlet as he actually did require would not have meant a loss of half the efficiency of the play-writing function! For Napoleon to have taken twenty instead of five minutes to plan a series of moves at Austerlitz would not have meant that his generalship was only one fourth as efficient!

The zero point of efficiency in the function of mental multiplication would be 'just not to multiply a number like 3 by a number like 2 in, say, ten minutes.' We do not of course know just at what point between this zero and the ability to multiply a four-place by a four-place number mentally in five minutes with only two figures in the answer wrong (as Miss Arai did at the beginning of work), we should place her ability, at the end of work, to multiply a four-place by a four-place number in eleven minutes. The reader may judge for himself. It is my impression that she could, at the end of work, certainly have multiplied a three-place by a three-place number (and probably a four-place by a three-place number) as quickly and accurately as she had multiplied a four-place by a four-place number at the beginning, and that it would be absurd to place the efficiency of her last half hour's work in each period at less than 75 per cent of the initial efficiency.

THE AMOUNT AND RATE OF FATIGUE OF A SINGLE FUNCTION

I shall present the results of such studies of the influence of continued exercise of a function for an hour or more upon its efficiency as can be made clear without elaborate explanation. The selection will not be prejudiced in favor of any view of the amount of fatigue; and the conclusions, based on the findings of many independent investigators, will be trustworthy. This security is the greater since, although most of these investigators expected continued maximal work for one or two hours to decrease efficiency to a large extent, their results almost unanimously show its effect to be very slight.

The functions in question involve purely intellectual work complicated more or less by sensory and muscular work. The question of distributing each decrease in efficiency amongst these three factors will not, however, be troublesome since there will be only very slight decreases, or none at all, to be distributed.

Burgerstein ['91] found, with a large group of children from nine and a half to fifteen and a half years old, in an hour's continuous work in adding and multiplying,* the following results:†

First	15	min.	Number	of	operations	28267,	errors‡	851
Second	15	"	"	"	"	32477	"	1292
Third	15	"	"	"	"	35443	"	2011
Fourth	15	"	"	"	"	39450	"	2360

If we assume that an error is equivalent to five operations, in the sense that on the average a child, by working more rapidly and doing five more operations than he did, would make one more error, we have more comparable measures, namely:

First	15	min.	35812	operations	with	2360	errors
Second	15	"	37817	"	"	2360	"
Third	15	"	37188	"	"	2360	"
Fourth	15	"	39450	"	"	2360	"

According to just what relative weight we thus allow to speed and accuracy in determining total efficiency, the later work will seem a little better or a little worse than the earlier. The work does about as much good by its training as it does harm by its continuity.

Höpfner ['94] measured the accuracy of a class of forty-

* Using such examples as:

     28704516938276546397
plus 35869427359163827263

and 28704516938276546397 × 2 (or 3, or 4, or 5, or 6, but not 7, 8, or 9).

† Combined from Burgerstein ['91, pp. 608 and 613].

‡ An error means a wrong figure in an answer.

six children in writing nineteen sentences from dictation in an examination for promotion. Because of the presence of two very deaf children for whom the sentences had to be repeated, two hours were given to the total test, or over six minutes per sentence of six or so words (30 letters). The experiment was thus not properly a measurement of the effect of continuous work upon efficiency, since the pupils were not pressed to their utmost, but of the varying quality of a very slowly given dictation exercise. Since, however, it is commonly quoted as a measure of fatigue, its results are given here.

Of the letters in the first 4 sentences	99%	were right,	1%	wrong
" " " " " next 4 "	98%	" "	2%	"
" " " " " " 4 "	97%	" "	3%	"
" " " " " " 4 "	96%	" "	4%	"
" " " " " last 3 "	95%	" "	5%	"

Holmes ['95] had school children, of grades 3 to 8 inclusive, add examples like $\frac{26753752624}{18932684395}$ for 36 minutes, divided into four nine-minute periods with 4 minutes rest between each pair. The answers were also all copied. The results were as follows:

	Errors			Amount	
	Adding	Copying	Both	Single Additions	Figures Copied
1st 9 min.	346	107	453	23713	23669
2nd 9 min.	430	115	545	27741	27773
3d 9 min.	643	208	851	29809	29808
4th 9 min.	812	182	994	20985	31258

Deducting five from the number of additions for each error in adding, and ten from the number of figures copied for each error in copying, we have as efficiency scores:

	Single additions, discounted for errors	Figures copied, discounted for errors
1st 9 min.	21983	22599
2nd 9 min.	25591	26623
3d 9 min.	25594	27728
4th 9 min.	26925	29438

Bettman and Aschaffenburg [Bettman, '95, p. 157] found, in a series of 1000 reactions with choice, average times for successive 200's of 323, 328, 336, 355 and 349. Cattell and Dolley report that simple reaction-time "is not greatly affected by . . . the number of reactions already made." ['95, p. 394]

Oehrn ['95] reports the efficiency during each 15 minutes of a continuous work-period of two hours for each of ten educated adults in the case of each of the following functions:*

Counting letters (in a Latin book) one at a time, "with the utmost possible speed, word after word and line after line. Whenever a hundred letters were counted, a pencil-mark was made at the appropriate place and counting began with one again." [Oehrn, '95, p. 96]

Counting letters three at a time.

Memorizing nonsense syllables, in series of twelve, made into four nonsense-words. The worker read such a series "from beginning to end until it could be reproduced perfectly once." [*ibid.*, p .100]

Memorizing lists of 12 digits.

Adding columns of one-place numbers, without writing the sums.

Writing from dictation, the one who dictated being careful to push the writer to his utmost accomplishment.

Reading at the fastest rate consistent with complete enunciation.†

Oehrn made no measurements to separate the possible loss in efficiency due to the continuousness of the work from the possible gain due to the practice which it afforded. He did not, that is, test his subjects after a rest. His figures

* Except memorizing nonsense syllables, to which only an hour and a half was given.

† For details concerning the administration of these experiments, see Oehrn, '95, pp. 94-107.

represent only gross achievement for each of the eight successive quarter-hours. They are given in Table 2. Each entry of the table gives the proportion in thousandths which the efficiency for the given period was of the greatest efficiency shown in any one of the eight periods.

Oehrn's results show that, in general, whatever fatigue there may have been was so slight as to be counterbalanced by the gain from practice including the adaptation or 'warming up' to the work. FIG. 2 shows the central tendency of the change in efficiency considering all six sorts of work together. In general that is, the subjects worked equally well throughout the entire two hours. This general result

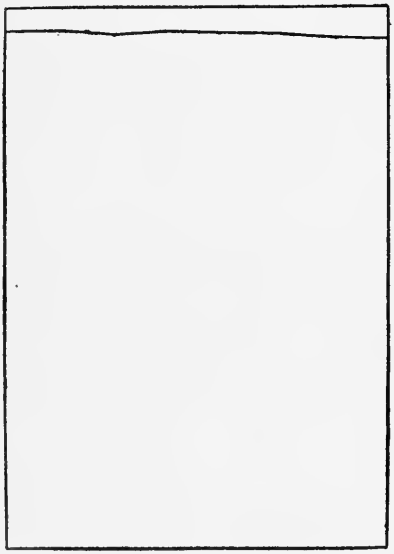


FIG. 2.

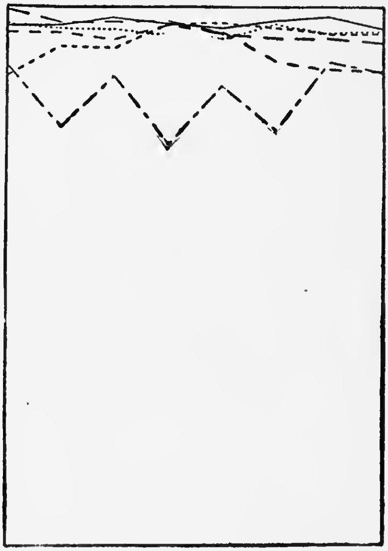


FIG. 3.

might have come from a steady improvement in some of the functions, balanced by a steady loss in efficiency in others, or from various rates of change in efficiency in the different functions. As a matter of fact, however, as FIG. 3 shows, the different particular functions follow closely the general tendency. Their slight divergences therefrom are probably due to the small number of subjects and experiments.

Each person, too, might have had a different work-curve, some getting better at points when others were getting worse. As a matter of fact, however, the ten individuals' work-curves follow closely the general work-curve which repre-

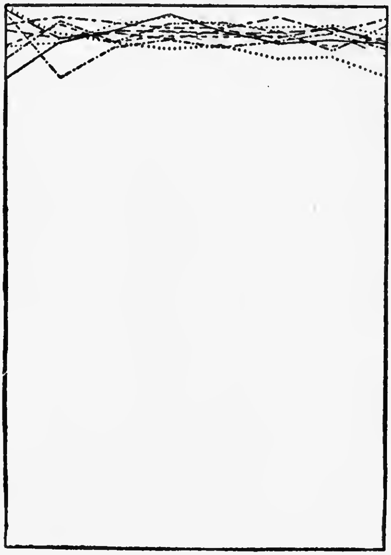


FIG. 4.

FIGS. 2, 3, and 4. The Results of Oehrn's Measurements of the Efficiency of Eight Successive 15-Minute Periods in the Case of Ten Subjects, Each Tested in Six Different Functions. FIG. 2 shows the central tendency of all sixty work curves. FIG. 3 shows the work curves in the cases of the six functions tested. FIG. 4 shows the work curves in the case of the ten subjects.

sents the central tendency of them all.* The facts are shown in FIG. 4.

In learning series of nonsense syllables during an hour and a half, Oehrn's subjects *did* lose in efficiency as the work progressed. The relations of the numbers of nonsense

* The slight divergences of the individual work curves from the horizontal are probably in part real. Individuals doubtless differ in the rate of change in efficiency with continuous work, as they do in almost everything. But there are no signs of any sharp divisions into 'quick adapters' and 'slow adapters,' 'early fatiguing' and 'late fatiguing,' or the like.

syllables learned in the successive 15-minute periods were as:

1000, 855, 826, 754, 666, and 751.

The meaning of the results obtained in memorizing nonsense-syllables is doubtful; for, as one learns more and more

TABLE 2.

SCORE OF SUCCESSIVE QUARTER-HOURS OF CONTINUOUS WORDS, TRANSMUTED INTO MULTIPLES OF THE HIGHEST SCORE OBTAINED BY THE INDIVIDUAL IN QUESTION IN THE FUNCTION IN QUESTION. After Oehrn, '95. (The order from left to right is the order of time.)

			Counting letters, one at a time.					
K.	0.849	0.875	0.935	1.000	0.957	0.971	0.942	0.978
H. D.	0.965	0.977	0.973	0.973	0.965	1.000	0.958	0.976
M.	0.970	0.962	0.971	0.939	1.000	0.967	0.995	0.956
B.	1.000	0.980	0.987	0.946	0.877	0.998	0.998	0.967
F.	0.901	0.950	0.991	0.960	1.000	0.988	0.986	0.949
W.	0.960	0.980	1.000	0.995	0.973	0.985	0.966	0.985
H.	0.777	0.826	0.915	0.950	0.920	0.931	1.000	0.945
E. D.	0.915	0.986	1.000	0.997	0.996	0.975	0.992	0.992
Frl. R.	1.000	0.948	0.973	0.928	0.944	0.873	0.950	0.954
O.	1.000	0.972	0.983	0.976	0.944	0.944	0.976	0.948
			Counting letters, three at a time.					
K.	0.878	0.934	0.958	1.000	0.978	0.935	0.937	0.922
H. D.	0.867	0.952	0.936	0.961	0.963	0.961	0.961	1.000
M.	0.881	0.919	0.965	1.000	0.977	0.925	0.859	0.921
B.	0.861	0.855	0.896	1.000	0.969	0.985	0.931	0.885
F.	0.837	0.912	0.906	0.976	0.984	0.963	1.000	0.950
W.	0.998	1.000	0.971	0.903	0.897	0.994	0.982	0.992
H.	0.822	0.861	0.879	0.911	0.970	0.973	0.966	1.000
E. D.	0.770	0.779	0.833	0.827	0.879	0.866	0.885	1.000
Frl. R.	1.000	0.941	0.911	0.919	0.906	0.935	0.874	0.891
O.	0.971	0.993	1.000	0.996	0.971	0.964	0.968	0.954
			Adding.					
K.	0.857	0.932	0.952	1.000	0.972	0.916	0.873	0.912
H. D.	0.889	1.000	0.923	0.897	0.908	0.840	0.885	0.874
M.	1.000	0.970	0.950	0.980	0.909	0.977	0.931	0.867
B.	0.983	1.000	0.945	0.953	0.942	0.899	0.884	0.930
F.	0.945	0.901	0.959	1.000	0.977	0.971	0.960	0.881
W.	0.896	0.895	0.872	0.913	1.000	0.900	0.866	0.795
H.	1.000	0.872	0.790	0.758	0.769	0.834	0.729	0.619
E. D.	0.912	0.990	0.907	1.000	0.967	0.857	0.889	0.935
Frl. R.	0.975	1.000	0.937	0.921	0.978	0.918	0.923	0.850
O.	0.961	0.941	0.977	1.000	0.928	0.894	0.879	0.889

TABLE 2 (Continued).

				Writing.				
K.	0.981	0.983	1.000	0.969	0.950	0.928	0.939	0.924
H. D.	1.000	0.982	0.970	0.939	0.935	0.932	0.908	0.927
M.	1 000	0.991	0.948	0.895	0.913	0.923	0.930	0.938
B.	1.000	0.958	0.962	0.982	0.952	0.956	0.958	0.936
F.	0.972	0.980	0.981	1.000	0.955	0.963	0.971	0.979
W.	0.976	0.981	0.967	1.000	0.956	0.952	0.954	0.973
H.	0.968	0.939	0.982	0.978	0.953	1.000	0.947	0.925
E. D.	0.990	0.954	0.984	0.985	1.000	0.997	0.974	0.969
Frl. R.	1.000	0.942	0.904	0.894	0.892	0.889	0.854	0.801
O.	1.000	0.964	0.976	0.925	0.918	0.930	0.905	0.894
				Reading.				
K.	0.964	0.936	0.985	0.952	0.949	1.000	0.985	0.946
H. D.	0.853	0.879	0.929	0.950	0.950	1.000	0.946	0.964
M.	0.940	0.885	0.910	0.885	0.934	0.978	0.991	1.000
B.	1.000	0.923	0.962	0.955	0.923	0.947	0.999	0.970
F.	1.000	0.998	0.974	0.987	0.915	0.949	0.915	0.943
W.	0.973	0.979	1.000	0.998	0.965	0.976	0.953	0.985
H.	1.000	0.960	0.928	0.928	0.933	0.941	0.894	0.887
E. D.	0.879	1.000	0.958	0.947	0.972	0.934	0.951	0.935
Frl. R.	0.965	0.962	0.927	0.894	0.956	1.000	0.997	0.975
O.	1.000	0.975	0.961	0.938	0.928	0.962	0.931	0.923
				Learning numbers.				
K.	0.454	0.535	0.545	0.575	0.848	0.767	0.595	1.000
H. D.	1.000	0.800	0.790	0.468	0.487	0.612	0.384	0.509
M.	0.778	0.842	0.914	0.938	0.985	0.785	1.000	0.833
B.	0.913	0.573	0.530	0.613	1.000	0.739	0.960	0.913
F.	0.554	0.529	0.831	0.614	0.858	0.692	0.418	1.000
W.	0.870	0.870	0.924	0.772	0.903	1.000	0.905	0.847
H.	1.000	0.839	0.977	0.952	0.907	0.907	0.954	0.921
E. D.	0.512	0.593	0.624	0.716	0.734	1.000	0.653	0.959
Frl. R.	1.000	0.950	0.941	0.961	0.676	0.735	0.891	0.785
O.	0.950	0.755	0.940	0.805	0.805	0.780	1.000	0.780

series, there develop interference and confusion. If *tib* has been connected with (1) *pon zet luf* (2) *nog, biz, ref,* and (3) *sib, kol, mek,* it thereby makes the learning of *tib, wif tek, saw,* or *tep, lin, tod, wak,* or *deb, nig, ron, puf,* and the like, harder. The first hour's learning may then reduce later ability by such interference and confusion, as well as by fatigue proper.

Bergström ['94] found that his efficiency in translating

German into English did not diminish in the course of four hours work, the average lines translated in twelve four-hour periods being, for successive half-hours, 24, 25, 30, 26, 28, 24, 28 and 29.

I calculate from Weygandt's data ['97] that twenty-four hours of rest increased the number of additions done by his three subjects in fifteen minutes by a fourth over the number done in the last fifteen minutes of a ninety-minute work period. This investigation, that of Vogt and that of Bolton, ['02] are the three cases of a substantial diminution of efficiency from exercise of a function without rest which the literature of fatigue affords.

Vogt ['99], who worked daily for ninety minutes at addition (one-place numbers, each to the sum so far obtained,

TABLE 3.

FATIGUE IN ADDING IN THE CASE OF VOGT.

Number of Additions done.

Date Mo.	Day	In the first 5 minutes of the work period	In the last 5 minutes of the work period
V	8	323	265
V	9	353	288
V	10	381	299
V	11	386	284
V	12	383	285
V	13	412	295
V	15	407	305
V	16	423	343
V	17	439	335
V	18	427	315
V	19	439	368
V	20	452	360
VI	1	380	293
VI	2	407	313
VI	3	455	343
VI	4	440	359
VI	5	483	402
VI	6	495	382
VI	7	502	397
VI	8	536	385
		Sum, omitting first day 8200	Sum, omitting last day 6231

till 100 was passed, a new series being then begun; nothing being written), with two rests of one minute each interspersed, and with distractions of various sorts to be overcome during the second half-hour, reports the results shown in Table 3. The amount done at the end of the ninety minutes was on the whole 76 per cent of the amount done at the beginning of work after rest the next day.

Vogt ['99] worked seventy-five minutes daily for twelve days at memorizing series of twelve digits (against additional distracting work during the second half-hour). The results for the first and last five minutes of work give an average loss from the beginning to the end of the same day, of 7½ per cent; and a gain at the beginning of one day over the end of the preceding day, of 9 per cent.

Vogt's results in the case of memorizing nonsense syllables showed a falling off of 27 per cent in the last five minutes of a seventy-five-minute period (of which two-fifths was spent in work against distraction) compared with the first five minutes ['99, p. 122]. When the work of the last five minutes of each day is compared with that of the first five minutes of the day following, the drop is 30 per cent. As has been noted, the interference and confusion resulting from the likeness of different nonsense syllables may have a part in this.

Continuous mental work of from three to eight hours failed, in the case of Dr. R. S. Woodworth, to produce any demonstrable diminution in efficiency. I quote the essential facts.*

Experiment 1.—The mental work done was to mark every word containing both *e* and *t* in 151 pages of a book. At the end of each minute a signal was given by a bell and the subject made a mark denoting the point he had reached. The work continued eight hours (from 10:15 A. M. to 6:20 P. M.) with only five interruptions amounting in all to less than nine minutes. The subject had done a very large amount of such work in connection with certain other experiments some

* From Thorndike, '00, pp. 573-576, *passim*, with some verbal changes.

months before, and found in a preliminary test of fifty minutes that he was nearly at a dead level (the amounts done in the several ten minutes being 238, 225, 231, 236 and 235 lines).

The lack of change in the amount done during the eight hours' constant work is shown in FIG. 5. The height of the

FIG. 5. The Curve of Work for Woodworth during Eight Hours Spent in Marking Words Containing *e* and *t*.

line represents the subject's efficiency as far as concerns amount of work done. . . .

The change in the quality of the work done has not been determined for the whole eight hours. In the first twenty-five minutes of work there were 37 omissions, *i. e.*, 6.075 omissions per hundred lines; while during the last twenty-five minutes' work there were 48 omissions, *i. e.*, 7.92 omissions per hundred lines. Thus there were 30 per cent more omissions at the end. This does not mean that the quality was 30 per cent worse, for six omissions in a hundred lines means six omissions out of about 200 cases. If we give the per cents for the number correctly marked we have only nine-tenths of one per cent difference between the earliest and latest work. It is not possible to tell just how much inferiority in work is represented by an extra omission per hundred lines.

Experiment 2.—In three hours' work in estimating the areas of small parallelograms of paper the accuracy of W.'s judgment was constant for the first two hours, but fell off 7 per cent in the last hour.

Experiment 3.—W., who had had some special preliminary drill in memorizing numbers, worked constantly from 3 to 7 P. M., memorizing sets of numbers written on a series of cards. Fatigue, if present, did not counterbalance the practice effect shown by FIGS. 6 *a* and 6 *b*, which represent the change in the quantity (time required) and quality (number

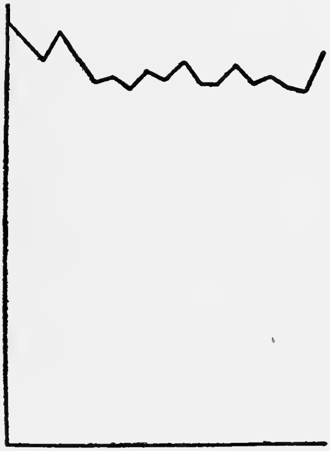


FIG. 6-*a*

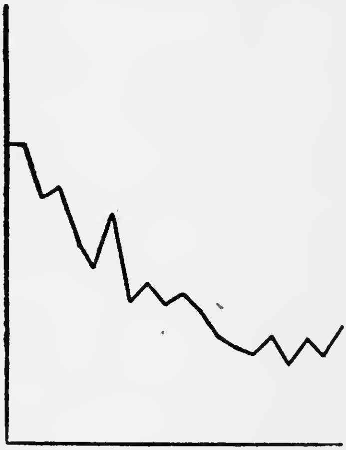


FIG. 6-*b*

FIG. 6. The Curve of Work for Woodworth during Four Hours Spent in Memorizing Numbers. *a* is the curve for time required; *b* is the curve for the mistakes made.

of mistakes) of the work. The slight rise in the curves at the end is probably the result of a growing dimness of light noted by the subject at the time of the experiment as troublesome. Apart from this there is no sign of any inability to work.

Experiment 5 (Experiment 4 is not reported here).—Another experiment was taken directly from ordinary life. W. had to go over three hundred and fifty cards on which were written titles of foreign books or articles and decide in each case whether to insert it in a certain bibliography. The work involved careful reading and translation, ability to remember the meaning of technical terms, decision as to the fitness of the article and finally decision as to how to classify it. The work took three hours with thirty-five stops of from ten to fifteen seconds (to rest the subject's eyes). The time taken to do each successive ten cards was recorded. No fatigue effect was observable. Fig. 7 shows the changes in the time required in the course of the experiment.

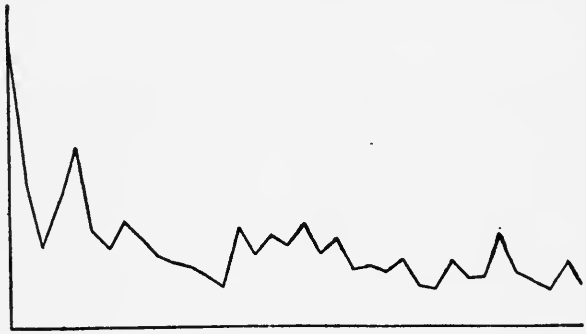


Fig. 7. The Curve of Work for Woodworth in Classifying Titles during Three Hours. The heights of successive points in the curve represent the times required for successive sets of ten titles

Lindley's ['00] three subjects gave as a median result for continuous adding for one hour the curve shown in Fig. 8, the fatigue effect of that much continuous exercise being fully counterbalanced by its practice effect.

Bolton ['02, pp. 200 ff. and 226 ff.] found the number of additions in successive fifteen-minute periods of an hour to be for one subject as 100, 86, 82 and 75. In the case of two-hour periods with the same subject they were as 100, 86, 83, 81, 78, 75, 76 and 75.

Kafemann ['02] worked at adding for ninety minutes on each of eight successive days, two five-minute rests being interspersed in each period. The work of the last five minutes was only two per cent less than that of the first five minutes of the next day after full rest.

Heüman ['04], who had six subjects add for sixty minutes (as in column addition, not passing 100), each one on a number of days, found the following proportions (Table 4)

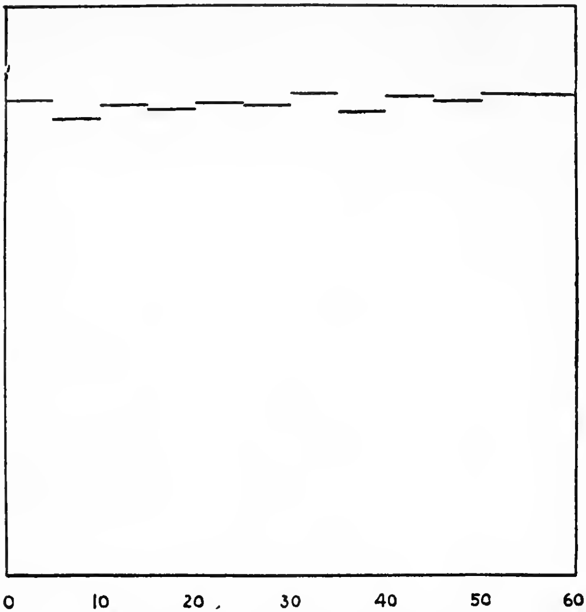


FIG. 8. The Median Curve of Work for Lindley's Three Subjects in Adding During One Hour.

holding amongst the amounts done in successive five-minute periods, the work of the first five minutes being taken as 100 in each case:

TABLE 4

FATIGUE IN ADDING. After Hëuman. '04.

Relative amounts done in each successive five minutes of an hour: Each entry is an average of several experiments' results.

	Period											
Individual	1	2	3	4	5	6	7	8	9	10	11	12
P.	100	93.6	93.9	97.6	101.9	101.4	101.5	106.5	100.4	104.4	105.6	106.6
C.	100	97.1	102.3	102.5	105.4	109.1	109.8	111.9	105.4	108.0	112.3	109.6
Eg.	100	99.4	101.7	102.3	104.8	103.7	103.9	105.6	106.3	104.8	103.4	102.5
H.	100	98.3	97.8	98.9	98.7	98.2	98.9	97.8	95.6	94.7	93.7	95.2
T.	100	98.1	99.8	101.0	101.7	101.9	101.4	101.7	100.3	100.4	97.7	96.9
W.	100	99.4	100.8	102.3	101.5	102.2	102.6	101.6	102.1	100.6	100.2	98.4
Average	100	97.7	101.1	100.8	102.3	102.8	103.0	104.2	101.9	102.2	102.2	101.5

Specht ['04] finds pronounced fatigue in the course of even the first ten minutes of single addition (written). A pause of five minutes between the first and last five minutes results, according to him, in a marked improvement of the record. The fatigue must in such cases be due, it would seem, to the writing at top speed rather than to the purely mental factors. Specht's results are so contrary to general experience in mental work, and even to general experience of the curve of this particular form of work, that they demand verification by other students.

Whipple ['10, pp. 262 and 265] reports that in approximately two hours' continuous work in marking the letter *c* or *o* on pages of mixed letters, the achievements in successive half-hours were as 93.6, 88.2, 86.3, and 92.1.

The author ['11] secured records of continuous work, or work broken only by a brief rest for luncheon, of from four to twelve hours from fourteen adult students. They also worked for an hour or so on a later day after rest. The work was the mental multiplication of a three-place by a three-place number. The amount of fatigue may be inferred from the difference between the time required to do the same amount of work with the same accuracy at the end of the work-period and at the beginning of the test after long rest.

The facts, including two individuals whose work-periods were less than four hours, are given in Table 5. Table 5 (A) gives the gross scores—that is, time taken and errors made—at the end of work and at the beginning of the trial on a later day after rest. Table 5 (B) gives the scores, each reduced to a single measure of efficiency by adding to the time taken for each example 12 per cent for each wrong figure in the answer and for each answer that was wrong.*

* Thus, if an example was done in 200 seconds with 1 error, the efficiency score would be 200 + (.12 × 200) + (.12 × 200). If the gross score was 200 seconds with 3 errors, the efficiency score would be 200 + (.36 × 200) + (.12 × 200).

TABLE 5.

FATIGUE IN MENTAL MULTIPLICATION WITH THREE-PLACE NUMBERS.

			A. Gross Scores									B. Scores expressed as time required, accuracy being equal			
			Next to last 4 examples of work-period			First 4 examples after rest			Second 4 examples after rest						
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Individual	Length of work-period in minutes	Length of lunch pause in minutes	Time in seconds	Digits wrong	Examples wrong	Time in seconds	Digits wrong	Examples wrong	Time in seconds	Digits wrong	Examples wrong	Next to last 4 examples of work period	First 4 examples after rest	Second 4 examples after rest	Increase of time required, with equal accuracy, at the end of work-period over that at the beginning of rest-period. In percents of the latter $\frac{13 - \frac{1}{2}(14+15)}{\frac{1}{2}(14+15)}$
1	467	0	1230	1	1	850	1	1	790	2	2	1302	925	870	45
2	428	0	375	0	0	240	2	1	275	1	1	375	260	290	36
3	300	0	840	1	1	360	0	0	240	0	0	890	360	240	197
4	262	0	1540	5	3	910	2	2	915	5	2	1920	1020	1105	81
5	212	0	600	10	4	660	12	4	630	11	4	852	980	900	— 9
6	121	0	760	13	4	780	13	4	1050	11	4	1140	1120	1525	— 15
7	734	29	2520	8	3	2100	8	3	2460	3	3	3360	2790	3120	14
8	490	77	505	6	3	460	2	2	425	2	2	635	515	475	28
9	460	60	1235	11	4	780	7	4	960	2	2	1780	1040	1075	68
10	432	204	280	5	3	225	2	1	255	1	1	348	241	270	36
11	420	90	850	2	2	520	5	3	485	4	3	955	645	590	55
12	392	45	1645	4	3	1010	1	1	1115	0	0	1990	1070	1115	83
13	388	52	760	4	2	330	4	2	320	1	1	920	390	340	152
14	314	98	780	7	3	480	1	1	420	0	0	1015	510	420	118
15	240	24	720	2	1	465	3	2	885	2	1	786	536	965	5
16	240	30	170	1	1	295	1	1	255	1	1	177	313	270	— 39

It would be unwise to try to combine all these sixteen records into one average or median amount of fatigue. In the case of the seven individuals who worked approximately seven hours, the average increase in time required (with allowance for accuracy) at the end of the work-period over the beginning of the test next day was 54 per cent. For six individuals who worked from three and a half to five and a quarter hours, the average increase is 59 per cent. For the two individuals working eight and twelve hours it is 21 per cent. The corresponding medians are 55 per cent, 43 per cent, and 21 per cent.

In considering the percentages showing the amount of fatigue, the reader will bear in mind the significance of time as a measure of efficiency in this function, as noted in connection with Miss Arai's experiment. For a person to be able to multiply mentally a three-place by a three-place number *at all* means a rather high degree of efficiency in the function. A great diminution in efficiency should make a person unable to remember one partial product while getting another. For one to take twice as long means to be half as efficient only in one special sense of "half as efficient."

The *degree* of difficulty that can or cannot be met is for many purposes a more satisfactory measure than the time required and the success attained in meeting *one given* degree of difficulty. With mental multiplication it would be practicable, though very laborious, to make the former measurement, by having interspersed in and at the close of a rigidly fixed long work-period of mental multiplication with two three-place numbers, tests in multiplying a four-place by a three-place number, a four-place by a four-place, a five-place by a four-place, and a five-place by a five-place. There should also be tests with three-place by two-place, two-place by two-place, etc., in case an individual comes to fail utterly with the three-place by three-place multiplication.

The author ['12] found with five subjects who added columns, each of 10 one-place numbers, continuously for from one and a half to two hours per day, that the difference between the last ten minutes of one day's work and the first ten minutes of the work of the next day—that is, after the rest—was, for each individual in each pair of consecutive days, as shown in Table 6. The central tendency is toward a fatigue effect of 6% (i. e., to do 6% fewer examples in the same time) from approximately 100 minutes of continuous exercise at adding.

TABLE 6. FATIGUE IN ADDING

Time (corrected for errors) required to add *x* rows at the end of each work period and at the beginning of the following work period. *x* = 6 for *D*, 2 for *L*, 4 for *Mc*, 3 for *R* and 6 for *S*. Also the per cent which the latter is of the former in each case.

	1*e*. At End of Work Period 1	2*b*. At Beginning of Work Period 2	2*e*. At End of Work Period 2	3*b*. At Beginning of Work Period 3	3*e*. At End of Work Period 3	4*b*. At Beginning of Work Period 4	2*b*/1*e*	3*b*/2*e*	4*b*/3*e*	Av. *b*/Av. *e*
D	715	565	595	620	615	580	79	104	94	92
L	975	615	590	560	545	585	63	95	107	88
Mc	543	531	579	510	511	466	98	88	91	92
R	630	545	615	545	535	495	87	88	93	89
S	897	850	842	934	779	743	95	111	95	100

Average of all per cents = 93. Median of all per cents = 94.

Arai ['12] gives records of eleven adult subjects in continuous mental multiplication of a two-place by a two-place number (*e. g.*, 27 × 43, 68 × 54, 93 × 75) for two periods, each of two hours, separated by from one to forty-seven days (1, 3, 3, 7, 13, 14, 21, 26, 27, 34 and 47).

The score given is the time required per example, corrected roughly for errors by adding ten seconds for each wrong figure. The averages for the eleven subjects for each successive ten minutes are given in Table 7. Since the measure is the time required, the decrease in the size of the numbers means that there was an *increased* efficiency in spite of the two hours of continuous work. In the case of all but one of the eleven subjects, ten-minute tests of the same function were made at the close of one or both periods after a rest of ten, sixty or one hundred and eighty minutes. For the seven subjects thus tested after one to three hours rest,

TABLE 7.

FATIGUE IN MENTAL MULTIPLICATION WITH TWO-PLACE NUMBERS. After Arai, '12, p. 96.

Time taken per example (in seconds, plus ten seconds for each wrong figure) in each successive ten minutes of two hours continuous mental multiplication with two two-place numbers. Averages from 11 individuals:

	1	2	3	4	5	6	7	8	9	10	11	12
1st period	98.7	97.7	95.9	99.3	101.4	76.2	94.8	96.4	92.2	95.9	86.4	74.7
2nd period	69.9	65.5	70.4	74.1	67.3	76.5	68.5	79.8	72.4	65.3	73.2	69.4
Average	84.3	81.6	83.2	85.2	84.4	76.4	81.7	88.1	82.2	80.6	79.8	72.1

the time required per example (corrected for errors) in the last ten minutes of the work-period was on the average 112 per cent of that required after the rest. A ten-minutes rest, in two subjects who were tested with it and the longer rest, seemed as good as an hour's rest. If all ten subjects with all rests are used, the percentage of time required at the end of work to that after rest averages 121 (median, 112%).

The rest here has in part the effect of relieving the confusion due to the perseverance of harmful memories from previous examples, so that it may not be fair to assume that two hours of this work increases the time required by a tenth to a fifth. But, calling the increase fifteen per cent, and assuming that to take 75 seconds instead of 65 to multiply a two-place number by a two-place number means a decrease of absolute efficiency of 6 per cent, we still find this taxing and odious work reducing efficiency by only 3 per cent per hour. Probably 2 per cent or less would be a fairer estimate.*

The results quoted give a fair sampling of the amounts of difference found between the efficiency of a function after

* For the detailed facts from which the results of the last two paragraphs are computed see Arai, '12, Tables VIII, XX and XXI, on pp. 52-56, 95 and 96.

long exercise of it at the individual's supposed maximum exertion, and its efficiency after full rest. These differences are in general very slight. A man can work for several hours at his utmost, and at the end do nearly as well as he will after full rest. Except when the function exercised is very disagreeable, either in *toto* or in the degree of restraint which it demands, the loss during the work period is often indiscernible. Indeed there is usually a gross gain, though after full rest there is a further gain. Such statements as Binet's "Tout effort est accompagné d'une certaine fatigue" ['98, p. 302], give then a wrong impression of the amount and rate of fatigue.

Such statements have been common partly because those who have written about fatigue at all have usually been interested in it and ready to believe in its existence and magnify its amount, partly because the gross results found for fatigue of a muscle have misled expectations in the case of mental fatigue, and partly because many students of this problem begged the question by presupposing that mental work done without rest must decrease the ability to do further work, and so have been unduly ready to adopt methods where late tests were subject to loss in interest, and to choose as measures of decreased efficiency those features of the total function which did grow worse. Even Ellis and Shipe ['03], who found in a thorough study that the amount of work done made no difference in the efficiency of any of the functions which they tested, are still inclined to insist that there must have been a decrease in efficiency, the methods of testing it being at fault!

It may be that more complex, more creative, and more important functions—such as writing poetry, deciding a nation's policy, or developing a scientific theory—will be found to suffer more from lack of rest than the more mechanical operations that have been measured. There may be a gradation such that when one can no longer write an epic, he can still write an elegy, and that the loss of the latter power may still leave ability to write 'verses to be read at a dinner,' and so on through writing a news-letter, writing directions to a servant, down to writing one's name. I doubt, however, whether complexity, originality and importance in and of themselves make a function fatigue much more rapidly. Tests in completing couplets of verse, solving originals in geometry, and working with large administrative tasks would not, in my opinion, show very great differences from the results already quoted for simpler functions, if excitement and worry were eliminated and if the interest in the work were kept at the same level. Such experiments with complex and creative functions are very much needed.

To the reader the fact that a man works no better after two hours of work and two hours of rest than he did at the end of the two hours of work will appear to be sufficient evidence that the work caused no deficiency which the rest cured. It *is* sufficient evidence, at least for the case of any of the functions with which we have been concerned in the workers with whose scores we have been concerned. Its sufficiency has been doubted, however. Men, being convinced that there *must* be a large temporary diminution in efficiency due to work and curable by rest, but finding little or no such loss from beginning to end of work, and little or no such gain from beginning to end of rest, have hit upon a peculiarly subtle and plausible means of inferring that, after all, it was really there.

They make the supposition that *any decrease in the permanent improvement* (resulting from, say two hours of continuous work) *below the permanent improvement that would have been made if the two hours had been distributed in the best possible manner,* is to be reckoned as fatigue. This argument sounds innocent, but is essentially unsound. It confuses a temporary deficiency which rest can cure with a permanent deficiency which *rest cannot cure,* but which a better distribution of the practice periods *could have prevented.* For example, suppose that I had argued concerning the five subjects mentioned on page 37f. who added for two hours daily for four days, as follows: An hour of exercise in adding 10-digit examples, distributed in seven daily work-periods of approximately equal length, produces in educated adults a reduction in the time (corrected for errors) of 30 per cent.* Two hours of exercise distributed over 14 days may fairly be expected to produce an improvement of at least 38 per cent. Two hours of continuous practice, however, produced in five educated adults a change of less than 8 per cent improvement. Therefore the fatigue due to the two hours work is at least 30 per cent.

Every one of these statements is true except the last. Its falsity is revealed by the following facts: First, on the next day at the beginning of work these subjects did not do 30 per cent better. They did, as a matter of fact, 12 per cent better. Second, if the experiment is continued, and fatigue is again estimated by this method of comparison of actual gain or loss with the gain from equal exercise optimally distributed, the amount of fatigue decreases enormously. No one acquainted with the facts would claim for the last two hours of eight optimally distributed, a practice gain in educated adults of over 13 per cent. It would probably be very, very much less. Now in two continuous hours forming the seventh and eighth in the experiment, the general change was a loss of 1 per cent. The fatiguing effect of this two hours is thus, at the most, 14 per cent, or half of what it was a few days earlier. It much more probably is 7 per cent, or only a fourth of what it was a few days earlier. By those who estimate fatigue by the amount of difference between (a) the gain that would have been made with an optimal distribution of practice periods and (b) the gain that was actually made when no rests were taken, this discrepancy has to be explained by supposing that the four two-hour periods of exercise have

* This has been found to be substantially correct by Thorndike and by Wells.

had a three-fold effect: (1) that they have permanently improved ability as much as eight hours optimally distributed exercise would have done, but (2) they have also temporarily so diminished ability as to counterbalance this on each day and (3) they have also enormously diminished their power to temporarily diminish ability. Unless the first and third of these statements are true the second cannot be. The first we have seen to be false in the case of the two hours exercise.

It must *always* be false unless the optimal distribution of time is that of the work period of the fatigue experiment in question. For one distribution of time, D_1, to give the same amount of permanent improvement as another, D_2, can mean only that in later trials under identical conditions those who had been subjected to D_1 show the same superiority to their earlier scores as do those who have been subjected to D_2.

In general, if two random groups, *A* and *B*, of equal ability at the date D_0 in the function *M*, are both given *X* hours practice during the time from D_0 to D_n, and if in a selection of tests after D_n that are random as to date, length and all other external conditions, groups *A* and *B* show equal efficiency, then, no matter what difference there was between *A* and *B* in the way the *X* hours were distributed, that difference can have made no difference in the amount of permanent improvement in *M*. The measure of the permanent effects of continuous and broken exercise upon efficiency is not in some mythical changes which would show themselves if fatigue, warming up, additional practice and the like did not interfere and hide them. It is in a fair sampling of some actual operations of the function.

In short, permanent improvement can mean only the gains shown, first, at the end of work plus a length of time such as just provides full rest, and second, at any given selection of later dates. If continuous work does produce the same permanent improvement as optimally distributed work, then the decrease of the score at the end of a two-hour period below that at the beginning of the next is the measure of its smaller temporary improvement. If continuous work *does not* produce the same *permanent* improvement as optimally distributed work, the measure of its smaller *temporary* improvement cannot be the decrease in the gain from *X* hours' continuous work below what would have appeared if the *X* hours had been optimally distributed.

Exercise of a function without rest shows two radically different effects. One is that the function is, at the end of the exercise, slightly less efficient than it becomes after a certain amount of rest. The other is that it is still less efficient than it would have become if the exercise had been distributed optimally so as to prevent over-learning, loss of satisfyingness, practice at less than maximal effort, and the like. The two effects should not be confused. Just as it was shown, by the experiments with hunger and nasal obstruction, that a function's efficiency could be greatly lowered temporarily with but slight effect upon the permanent improvement from the practice, so it can be shown that a function's permanent improvement from certain practice can be much reduced, though the efficiency is lowered temporarily slightly or not at all.

The fallacy of arguing that long use causes a temporary deterioration because the permanent improvement is less with it than it would have been with the same time in short periods is much the same as in the argument that a man who had a thousand dollars at 10 A. M., January 1, and the same amount at 10 P. M. of the same day, must really have lost a fifth of it, since, had he speculated successfully with it on six or eight separate occasions during the day, he would have increased it to twelve hundred dollars.

CHAPTER III

The Curve of Work

In Arai's sample experiment, reported on page 14 ff., fatigue did not progress at the same rate throughout the work-period. This is shown in Fig. 9, which gives the average time required for each successive set of four examples. At a period around the fifth hour when the eighth and ninth sets were being done, efficiency did not decline. Again, at a period around the ninth hour when the thirteenth and fourteenth sets were being done, efficiency actually improved. In the sixth hour (from the ninth set to the tenth set of examples) there was a great falling off in efficiency. In the last two hours, the gain made in the ninth hour was all lost.* These apparent ups and downs in Miss Arai's curve of efficiency are illustrations of the fact that the decrease in efficiency as work continues may not be regular.

In general, it is obvious that the same amount of decrease and average rate might come as a result of any one of an indefinite number of work curves.

The available experimental material is inadequate for measuring changes in the rate of fatigue clearly and surely. For, unfortunately, the great majority of studies of continuous mental work have not included sufficient preliminary work to put the function beyond the point where a considerable discount has to be made for practice. Nor have they included

* The rapid diminution of efficiency from the ninth to the tenth set of examples, or toward the end of the sixth hour of the work-period, characterizes each of the single-day records and is probably characteristic of the individual. Arai interprets it and the later fluctuations as due to her habit of spending the time from 5 to 8 P. M. in recreation, of beginning mental work again at 8 and stopping at 10 P. M.

work-periods long enough to show the extreme influence of work done without rest. Nor have they equated 'amounts done' and 'errors made' into an intelligible measure of efficiency.* Nor have they included enough work-periods for

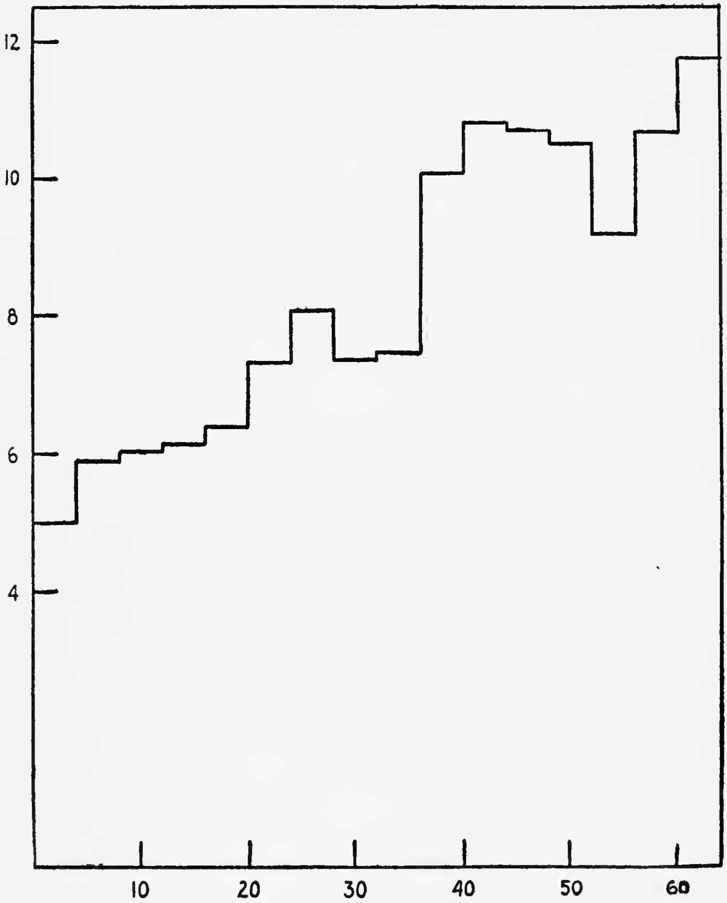


FIG. 9. Arai's Curve of Work in Mental Multiplication. The abscissa is scaled for the number of examples; the height of the curve over each division of the abscissa represents the time required (per example) for that set of four examples, with allowance for errors made.

* Kraepelin ['03, p. 21] excuses his failure to take account of errors as follows: "Apparently one can be content, in scoring the results, with the number of additions. The percentage of errors in so simple computation is usually very small, at least with adults."

each individual to allow us to decide with surety whether a given feature of the course of efficiency, for instance, 'warming-up' or 'end-spurt,' is truly characteristic of him in that function. Still less do they permit conclusions as to whether any given feature is characteristic of the individual in all functions, or of all individuals in that function. They do, however, make certain negative conclusions highly probable; and they need to be reviewed because of the prevalence of various theories about the curve of work—theories which have never been fully tested by even the inadequate facts at hand.

I shall report the results of one set of experiments in some detail as a means of making clear how each of the supposed irregularities of the work-curve is to be detected, and give only the essential facts in the case of the others.*

Various notions have been entertained concerning changes in the rate of decrease in efficiency of a function in the course of continuous exercise,—that is, changes in the slope of the curve of work. For instance, we have the doctrine that at the very beginning a person tends, other things being equal, to work at a higher efficiency than ever again. This we may call the doctrine of the 'Initial Spurt.' A second doctrine is that knowledge that the end is at hand produces in the last fraction of a work-period of limited length, other things being equal, a marked rise in efficiency. This we may call the doctrine of the 'End Spurt.' A third doctrine is that for the first half-hour or so, other things being equal, efficiency increases gradually. This we may call the doctrine of 'Incitement' or the 'Warming-Up' effect. A fourth doctrine is that a slower, longer and more lasting gain exists alongside of the 'Warming-Up' effect, called Adaptation. A fifth doctrine is that any more than usually rapid decrease in efficiency, by attracting the person's attention and rousing him to greater exertion, tends, other things being equal, to be followed by a relative increase in efficiency and maintenance of efficiency at a relatively high level for a few minutes.

* Much of what follows is quoted from an article by the author in the *Psychological Review*, vol. 19, pp. 165-194.

This we may call the doctrine of 'Spurt after Fatigue. A sixth doctrine is that slight ups and downs in efficiency come rhythmically in correspondence with fluctuations of attention, each total 'wave' being about two seconds long.

It should be noted that the terms Initial Spurt, End Spurt, 'Spurt after Fatigue,' 'Spurt after Disturbance,' 'Warming-Up' (Anregung), 'Adaptation' (Gewöhnung), 'Rhythm of Attention,' and the like, may each be used in two meanings. They may refer to objective changes in the efficiency of the function,—that is, in the height of the curve—or to imagined causes of such objective changes. Thus, End-Spurt may mean either 'an increase in efficiency in the last five or ten minutes of work,' or 'a reinforcing potency from knowledge that the end is near.' 'Adaptation' or Gewöhnung may mean either 'a rise in efficiency, slower than the rise called Warming-up and less permanent than the rise due to practice,' or some real factor which causes this rise and is different from the factors causing 'Warming-Up' or the practice effect.

I shall in this chapter use these terms only in the former objective meanings of changes in the efficiency of the function, asking, for instance, in the case of Initial Spurt, "To what extent is a high degree of efficiency appearing in the first few minutes of work characteristic of work curves in general, or of certain individuals in certain kinds of work?"

INITIAL SPURT

This phenomenon is certainly not characteristic of work curves in general. In the case of the 37 work-periods of 16 subjects engaged in mental multiplication (of a three-place number by a three-place number)* there was no evidence of it. In the case of five adults working at addition† (each for four two-hour periods), there is no evidence of it.

* The results of which are summarized on pages 75-80 of the *Journal of Educational Psychology,* Vol. II.

† The examples were each of 10 one-place numbers, printed in a very clear type 3½ mm. high.

TABLE 8.

SAMPLE RECORD OF GROSS SCORES.

Time taken and errors made in successive rows of 16 ten-digit examples, on four days after approximately the same amount of general mental work and on a fifth day after long rest. Subject *Mc*.

Row	Beginning at 2 P. M., March 10		Beginning at 1:10 P. M., March 11		Beginning at 7:30 P. M. March 12		Beginning at 1 P. M., March 13		Beginning at 9 A. M. March 17	
	Sec.	Err	Sec.	Err.	Sec.	Err.	Sec.	Err.	Sec.	Err.
1	150	1	138		185	1	172		110	
2	145	1	138		171		120		105	
3	145	1	133		171		110	1	105	1
4	133		122		173	1	119		100	1
5	133		131		172	1	105		105	
6	140		120		125		112	1	108	
7	135	1	115		110		110	1	100	
8	150	1	128		116		112		115	
9	153	2	132	1	130		115		95	
10	142	1	113		111		103			
11	160	2	124		115		118			
12	150	1	126		110		111			
13	140	1	118	1	112	1	105			
14	132		133	1	120	1	120			
15	150	1	135		130	2	127	1		
16	135	1	127	1	113		122	1		
17	157	1	128	1	109		124	1		
18	146		120		135		112			
19	143		133		117	1	115	1		
20	142	1	127		122		95			
21	168	1	128	1	106		125			
22	156		120		112		120			
23	158	1	120		116		111			
24	148		140		117		108			
25	138		138		118		106			
26	148		140		114		110			
27	147	1	135	1	123		112	1		
28	138		115		128	1	103	1		
29	147	2	145		138		120			
30	145		125	1	128	1	115			
31	137	1	135		110		118			
32	127		125		120	1	110			
33	137		130		130	1	115			
34	135	1	122		120	3	110	1		
35	133		132		117		108			
36	138		150	2	130		120			
37	122		132	1	138		116			
38	138	1	128		136	1	110			
39	132	1	144	2	122		113			
40	134	1			127		113			
41	130				133		110			
42	140				135		135			
43	130				127		120			
44	138				127		107			
45	130	1			122		125			

Since the results of this second series of experiments will be used throughout this chapter to illustrate the problems of the work-curve, they may best be described now.

Educated adults, graduate students of psychology, worked at adding for 1½ or 2 hours as continuously as possible, recording the time at the completion of each row of 16 examples. 21 different rows, each of 16 examples, were used to avoid memorizing any answers or sequences. The experiment was repeated four times on different days and finally a test of 10 to 15 minutes length was made. Each of the four long series was made under substantially the same external

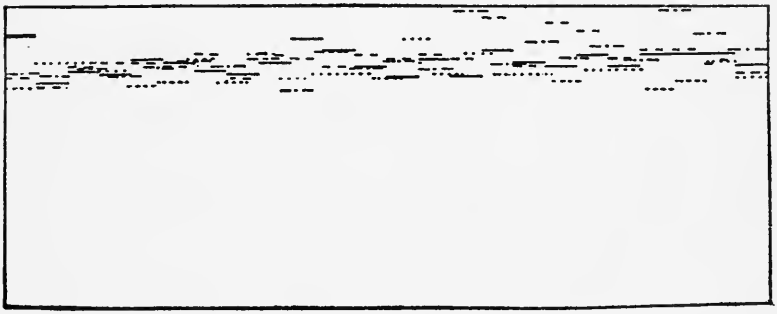


FIG. 10. Work Curves of Subject D in Adding: Four Days: Relative Time Required per Example in Successive Fractions of a Work Period of Approximately 100 Minutes. The long dashes all refer to the third day's work curves; the short dashes to the second day's; the dotted lines, to the fourth day's; and the dash and dot lines, to the first day's.

conditions; the fifth test was always made after a full night's rest or more. I append a sample record in Table 8. I have turned the gross scores into terms of a single variable by adding five seconds for each error.*

I shall present here the results from only five individuals. The results from the others are in agreement with every conclusion which will be stated.

I show in FIGS. 10 to 14 the work curves for the five subjects separately.

* I have also all the results reduced to terms of a single variable by a much heavier penalizing of errors, namely, an addition, per error, of one-fifth of the time for 16 examples. All the conclusions to be stated in this chapter are supported equally by the results by either method of equating speed and accuracy.

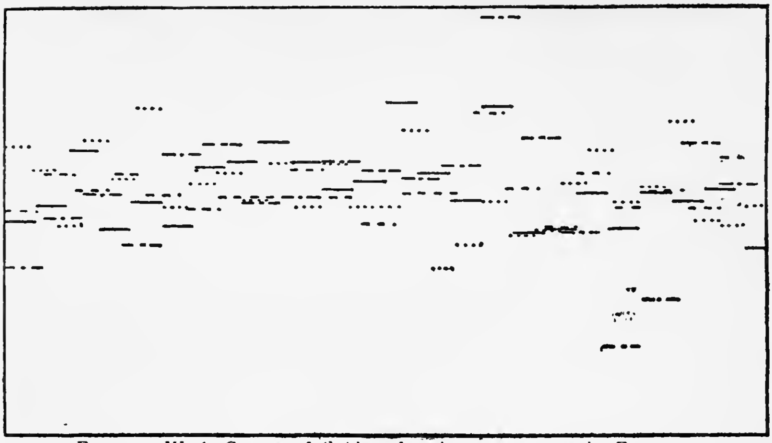


FIG. 11. Work Curves of Subject L. Arrangement as in FIG. 10.

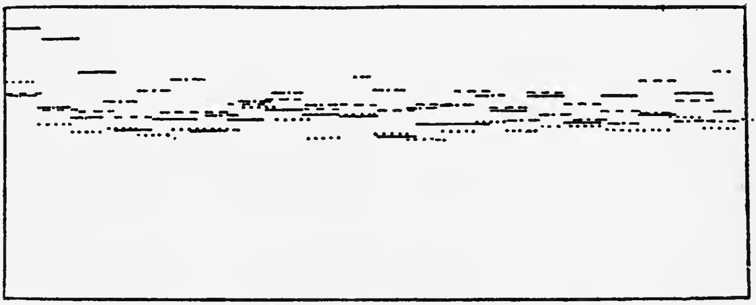


FIG. 12. Work Curves of Subject Mc. Arrangement as in FIG. 10.

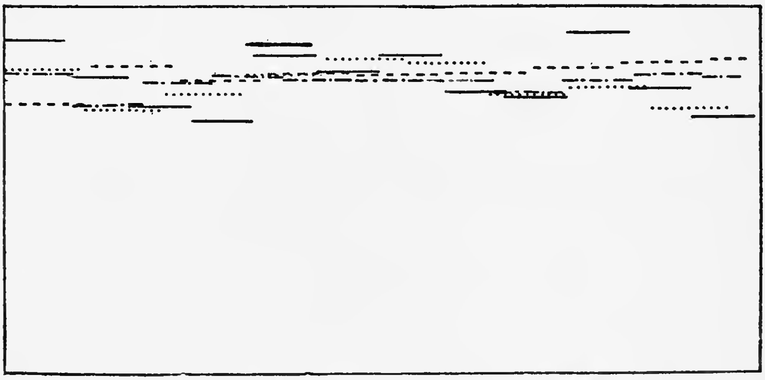


FIG. 13. Work Curves of Subject R. Arrangement as in FIG. 10.

Since efficiency is measured inversely by the corrected time per row, the lower the curve the greater the efficiency. Since one important fact to be shown by these curves is the variation of the same subject from day to day, the four curves from any one subject are so scaled vertically that the average achievement per unit of time is approximately equal on all four days, and are so scaled lengthwise that each successive fraction of the time is represented by the same abscissa-length on all four days.

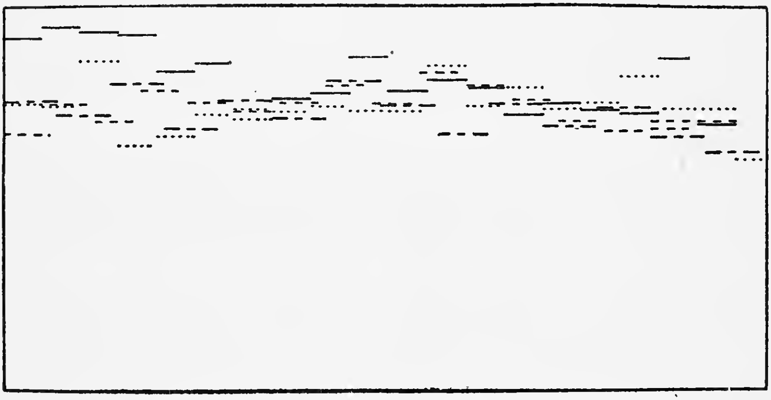


FIG. 14. Work Curves of Subject S. Arrangement as in FIG. 10.

I call the reader's attention particularly to the great variation in the form of the work curve for any individual from day to day.* Among the daily records can be found some which could be used, if taken singly, as excellent examples of initial spurt, end spurt, warming up, or fatigue, of various types. But such selection of a single day's record is obviously unjustifiable, since the same person on another day shows opposite fluctuations.† Many different days' records are

* The variation from day to day in the form of the curve of work may, in the case of one subject, *R,* be due in some considerable degree to the use of a too coarse unit in recording the time. But such is not the case with any other subject.

† The inconstancy of the same individual in respect to the form of the work curve would not have been so great if the work had been that of saying to oneself or writing down the results of *each successive*

necessary to determine the forms of an individual's work curve.

Initial spurt, if a real fact, will be found in an examination of the work, minute by minute, of the first quarter of an hour. Kraepelin supposes it to be a phenomenon chiefly of the first five minutes. It is possible to find two or three single records which might, if alone, be so interpreted, the most plausible case being one of subject *L's*: 265 seconds for the

TABLE 9.

DATA ON INITIAL SPURT.

Average time (corrected for errors) required to add successive rows of 16 ten-digit examples at the beginning of the four work periods. Five subjects—*B, L, Mc, R,* and *S.* Time in seconds.

Row	Subject				
	B	*L*	*Mc*	*R*	*S*
1	100	284	164	186	149
2	100	311	146	183	142
3	104	318	143	206	155
4	100	301	138	204	158
5	99		137	186	147
6	95		126	195	146
7	91		120		159
8	90		128		
9	113		136		
10	105				
11	96				
12	99				

Each 'row' equals 16 examples or 144 additions, involving the writing of 16 two-place numbers as answers, and one observation of the watch and record of the time.

addition, or if the subject had not expected a heavy penalty for an error, or if the subject had not pushed himself to the utmost, or if the subject had had a very great deal of special practice in adding at his maximal rate, but without trying to add by grouping. But in any case the variation in the form of the work curve on different days is so great as to require careful consideration of the unreliability of any determination based on only a few days' records. These unreliabilities have rarely been considered by Kraepelin and his followers, so that it is possible that the fluctuations to explain which they have invoked initial spurt, final spurt, spurt after fatigue, spurt after disturbance, or the rhythm of attention are one and all accidents of the unreliability of the determinations.

first 16 examples (no errors), 335 seconds for the next 16 (2 errors), and 360 seconds (4 errors) for the next 16. But these single instances are demonstrably not due to any consistent initial spurt. Exactly opposite instances occur, and in the same individuals. The average time (plus 5 seconds for each error) for each subject, for the four days' trials of

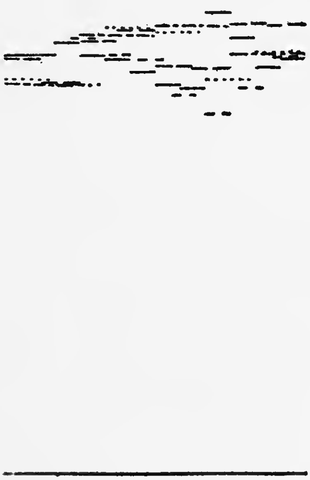


FIG. 15 FIG. 16

FIG. 15. The Absence of Initial Spurt. Average work curves of the first 20 minutes. Each sort of line (dot, short dash, longer dash, etc.) represents the average work curve of one subject. Heights equal times required to do equal amounts, with allowance for errors.

FIG. 16. End Spurts. Average work curves of the last 20 minutes. Each sort of line (dot, short dash, longer dash, etc.) represents the average work curve of one subject. Heights equal times required to do equal amounts, with allowance for errors.

each successive row is given in Table 9, including for each individual approximately his first twenty minutes' work in each of the four work periods.

The absence of any evidence of initial spurt in the facts of Table 9 is better realized by presenting them graphically. Let each centimeter horizontally approximate five minutes, let the same vertical height represent approximately the average efficiency of the function for each individual. We then have approximately FIG. 15, which shows the absence of any uniform fluctuation of any sort.

Or we may go back to the original scores and compute the work for each successive five minutes for each individual. The results are as in Tables 10 and 11. Weighting the results from all five subjects equally and averaging the five, we have the efficiencies for the four successive periods standing in the following proportions: 100, 99, 101 and 100.

TABLE 10.

DATA ON INITIAL SPURT.

Average number of examples done (each involving 9 additions, and the writing of a two-place number*) and of errors made, in each five-minute period of the first 20 minutes of work.

	D		*L*		*Mc*		*R*		*S*	
	Ex.	Er.	Ex.	Er.	Ex.	Er.	Ex.	Er.	Ex.	Er.
First 5 min.....	48.8	1.7	16.8	1.5	31.3	.7	26.3	.7	33.6	1.4
Second 5 min..	49.9	1.0	15.8	1.7	34.8	.8	24.5	1.2	31.6	1.5
Third 5 min....	48.2	1.0	15.6	2.0	37.8	.7	25.5	1.9	33.5	1.7
Fourth 5 min..	48.6	.8	16.2	.7	37.4	1.1	24.9	1.9	31.5	2.3

TABLE 11.

DATA ON INITIAL SPURT.

Per cent which the efficiency in each successive five-minute period of the first 20 minutes of work is of the individual's average efficiency for the 20 minutes.

	D	*L*	*Mc*	*R*	*S*
First 5 min....	100	104	89	104	103
Second 5 min..	102	98	99	97	97
Third 5 min....	98	97	107	101	103
Fourth 5 min..	99	100	106	98	97

I am unable to find anywhere any evidence of consistent initial spurt with any individual in all mental functions or with all (or nearly all) individuals in any kind of mental

* Because of the time required to write the answers and to record the time after every 16 examples, the number of additions that would have been made in mere addition may be taken to be not 9 but 10 times the numbers entered in the table.

work, much less with all individuals in all kinds of work. The work curves obtained by Oehrn, ['95] Amberg, ['95] Weygandt, ['97] Lindley, ['00] and other workers in Kraepelin's laboratory give no such evidence. Nor is it to be found in the data got by Yoakum, ['09] so far as he presents them.

Lindley's work, which was the most extensive, showed as speed ratios for successive five-minute periods at the beginning of work, 100, 98, 97, 97, and 96, using the data from all three subjects. The first five minutes, that is, differed from the second substantially only as the fourth from the fifth. Putting together Weygandt's series I find ratios of 100, 97 and 95½ for the first three five-minute periods. Hoch and Kraepelin ['95, p. 431 ff.] showed, on the whole, ratios of 100, 99, 98 and 94 for the first four five-minute periods. Miesemer ['02] showed ratios of 100, 96, 98, 97. In fact the very results of Rivers and Kraepelin, ['96] to explain which Initial Spurt was specially invoked, give as ratios (by five-minute periods) 100, 87, 99, 101, 102, 102. Clearly the fact in them to be explained is the 87 of the second five minutes rather than the 100 of the first five.

Of the three individuals for whom Book recorded the amount done in each minute of a ten-minute daily test in typewriting, one showed a rate in the first minute about 3 per cent above the average rate for the ten, counterbalanced by a rate about 3 per cent slower than the average in the second minute; one showed a rate below the average but about one per cent above the rate of the second minute; these both wrote by the sight method. The third, writing by the touch method, where the acquisition of the proper set and the relearning or freshening of associations are more important, showed a rate over ten per cent below the average, five per cent below the rate of the second minute and about one per cent below what the second minute would have shown apart from the general rise during the first five minutes. (See Book, '08, p. 105 and p. 126, Figs. VI and XI.) The average result would be toward an initial spurt of one per cent in minute number one followed by a sub-initial decline of one per cent in minute number two. In writing the same sentence over and over again, there was no initial spurt whatever [Book, '08, p. 114, Fig. IX]. The fact that so careful an investigator as Book still somehow extorts support for the doctrine of Initial Spurt from these facts shows the potency of an attractive term.

From the detailed scores given by Arai ['12] in Table IV [p. 38 f.], it appears that her work of the first five minutes at four-place mental multiplication was no more and no less efficient than her average work of the first fifteen minutes. The average time for the first example on the four days was 5.1 minutes; for the first, second and third examples, it was 5.2 minutes. The average number of wrong figures per answer was 1.25 for the first five minutes (i. e., for the first example on each day); and exactly the same for the first fifteen minutes.

For Arai's eleven adults who worked two hours at two-place mental multiplication, the work of the first ten minutes was a little *less* efficient than that of the second or third ten minutes (the average times required per example corrected for errors being, for these three successive sixths of an hour, 84.3, 81.6 and 83.2). This lower efficiency at the start is doubtless due to the general influence of practice, but, even with a liberal discount for practice, the figures give no evidence of initial spurt of any magnitude.

The one set of experiments which gives evidence of initial spurt is that of Vogt ['99] who, in a series of twenty days of addition for seventy-five minutes per day, did, in successive five-minute periods from the beginning of work, amounts standing in the proportions of 100, 94, 91½, 92, 86 and 85; and, in a series of ten days of memorizing nonsense syllables for the same time, did, in successive five-minute periods from the beginning, amounts standing in the proportions of 100, 79, 82, 78 and 73. In learning series of twelve digits, however, he showed just the opposite, the numbers learned in successive five-minute periods being in the proportions of 100, 106½, 106, 96 and 98.

It is, I admit, very likely that some individuals in some kinds of work tend to fall off rapidly from the too exacting standard which they set themselves at the beginning, just as some tend to rise rapidly above the standard which they cautiously try at the beginning. But these idiosyncrasies must not be misinterpreted as a general law. Until some hypothesis can be framed about initial fluctuations which will enable one to *prophesy* the form of the work curve, if only for a single individual or for a single function, these idiosyncrasies belong under the heading of 'accidental' or 'chance' variations —that is, of fluctuations whose causes are unknown.

END SPURT

It is often the case in ordinary mental work with a time limit, that as one approaches the end of the work period, the knowledge that he is approaching it leads him to spurt. In ordinary mental work one does not work throughout at one's possible maximum, so that such a spurt is easily possible. In experimental work, when the subject is required to work throughout at maximum possible efficiency, such a spurt can come only if the subject has deliberately disregarded instructions, or if the knowledge of the approaching end releases forces over which he had no control. The latter is apparently possible, various external stimuli, such as other competing individuals, applause and the like, being apparently able to add a reinforcement beyond what a subject's own determination can summon.

We may estimate* how common an occurrence such a final spurt is in work carried on with no deliberate reservation of

* For a demonstration of the existence and amount of end-spurt one should have subjects work to the end of a given period with and without knowledge that the end is approaching, all other conditions being identical.

power, and of how great a magnitude it is, from the following facts:

It is not common or important enough to make the last eighth of a two-hour period one per cent more efficient than the average eighth in the ten individuals each measured in six sorts of work by Oehrn.

In 16 subjects working for very long periods (4-12 hours) the effect of end-spurt in those who worked to a time limit is not great enough to counterbalance the effect (real or possible) of the decision of some to stop work because of having made a series of very inferior records. [Thorndike, '11]

In the four two-hour tests in addition previously described,

Table 12.

Data on End Spurt.

A.

Average time (corrected for errors) required to add successive rows of 16 ten-digit examples at the ends of the four work periods. Five subjects—*D, L, Mc, R, and S.* Time in seconds.

	D	*L*	*Mc*	*R*	*S*
Eleventh from last row added...	105				
Tenth from last row added......	104				
Ninth from last row added......	104				
Eighth from last row added.....	108		128		130
Seventh from last row added....	113		130		134
Sixth from last row added......	103		126		148
Fifth from last row added.......	110		125	205	142
Fourth from last row added.....	108		126	205	141
Third from last row added.....	110	297	143	203	137
Second from last row added.....	101	254	129	188	138
Next to last row added.........	104	365	130	190	110
Last row added....	95	278	134	200	126

B.

Average time (corrected for errors) required to add successive equal amounts (3 rows for *D,* 1 row for *L,* 2 rows for *Mc,* 1 row for *R,* 2 rows for *S*) at the ends of the four work periods.

	D	*L*	*Mc*	*R*	*S*
Fifth from last..........				205	
Fourth from last........				205	
Third from last	313	297	256	203	282
Second from last........	324	254	251	188	283
Next to last............	328	365	272	190	275
Last....................	300	278	264	200	236

not one subject of the five showed a *consistent* tendency to increase his efficiency by even five per cent during the last 5 or the last 10 minutes. For the five subjects whose records have been computed in detail the facts are given in Table 12 and FIG. 16.* There is a probable tendency to fluctuation in the direction of 5 per cent greater efficiency in the last few minutes, but the central tendency is almost certainly not to a change of over + 10 per cent, and is perhaps to as little change as 0.

Amberg's average results ['95] at the end of one- or two-hour periods of adding were, by five-minute periods, in the ratios 100, 101, 101 and 96. Weygandt's average results ['97] in the last twenty minutes of 90 minutes' continuous adding were in the following ratios:

Subject A	102	95	100	107
Subject B	100	99	99	100
Subject D	100	99	102	98
Average	101	98	100	102

Of Lindley's three subjects ['00] one did a bit more, one a bit less, and one almost exactly the same amount of addition per minute in the last 5 as in the average of the last 20 minutes. Combining the data for all three there is not a difference of one per cent. Miesemer's results ['02] for the last four five-minute periods of an hour's adding were in the proportions 101, 100, 99, 106.

In the last three five-minute sections of a seventy-five minute period of addition, Vogt ['99, computed from the data of his pages 92 and 93] did amounts standing in the proportions of 100, 94 and 92. In the last three five-minute sections of a seventy-five minute period of memorizing lists of 12 digits, however, ['99, computed from the data of page 121] he did amounts standing in the proportions of 100, 111 and 105; and in the last three five-minute sections of a seventy-five minute period of memorizing nonsense

* In FIG. 16 (p. 54), each centimeter horizontally approximates 5 minutes; the same vertical height is used to represent approximately the average efficiency of the function for each individual.

syllables (computed from his page 122) he did amounts standing in the proportions 100, 92 and 87. Averaging the three results, we have 100, 99 and 95, showing on the whole no evidence of end spurt in his forty-one work-periods.

Book's three learners of typewriting showed ['08, p. 105 and p. 126, Figs. VI and XI] in the last four minutes of ten-minute tests (enough days being used to give a very accurate record in each case) amounts done of approximately:

136.5, 132.5, 134 and 137.5, for subject X
170, 171.5, 176 and 175, for subject Y
119, 120.5, 118.5 and 122.5, for subject Z

On the average the amounts done in these four minutes stand in the relation 100, 100, 101 and 102.

In the case of Arai ['12] the last example (four-place mental multiplication) required 11.7 min., with 1.5 figures wrong (average of four days) while the last three examples required on an average 12.3 min., with 2.3 figures wrong in the answer. Allowing for the less accuracy, we may say that the amount done, at equal accuracy, would have been ten per cent greater in the last ten minutes than in the last half-hour.*

On the average, the eleven adults tested by Arai ['12, pp. 52-56] did two-place mental multiplication during the last ten minutes 5 per cent more rapidly than in the preceding ten minutes (of work periods of approximately two hours) but made 12½ per cent more errors per example. Allowing for the errors, we may estimate the gain in speed at equal accuracy as 3 or 4 per cent.

On the whole, no subject who has been tested four or more times shows consistently any great end-spurt; the general tendency is to a rise, in the last five or ten minutes, of three or four per cent in the amount of work per unit of time.

* Once again I remind the reader that one-tenth more done in this very difficult work does *not* mean one-tenth greater absolute efficiency.

SPURT AFTER FATIGUE AND SPURT AFTER DISTURBANCE

In mental work in ordinary life a person may obviously, if he is not doing his best, at any time do a little better to make up for an observed temporary deficiency, however caused. Deficiencies due to disturbances certainly, and to fatigue, if that acted unevenly throughout a work period, might be thus noticed and counterbalanced. In a subject who is keeping his efficiency at a maximum so far as he can control it, the observation of a fall in achievement might still so act as a reinforcement. Kraepelin and his followers assume that it does, and seem to regard each 'drop-rise' sequence in the curve of work as a deficiency caused by fatigue or disturbance which stimulates a gain in efficiency as a result of an Ermüdungsantrieb or Störungsantrieb.

It should be noted, however, that on general grounds the suggestion that one is doing well would seem more favorable to the efficiency of one already doing his best than the suggestion that he is doing badly, and that empirically no one has correlated the fluctuations in work curves with the incidence of disturbances of known character. The doctrine of spurt after fatigue and spurt after disturbance in the case of work done under the conditions of the ordinary fatigue experiment, is then at present a speculative hypothesis. It was devised apparently to explain the fluctuations in efficiency, from one minute or five-minute period to another, which are found in continuous adding, cancelling letters, memorizing, or other forms of mental work.

A rise following a fall in the curve easily attracts observation and tempts readily to theorizing. A rise followed by a rise or a fall followed by a fall is not so striking. The explanation of a 'fall-rise' sequence by spurt after disturbance or spurt after fatigue is really unwise, however. For if the fall is caused by a disturbance, no cause is required for the rise save the ending of the disturbance; while if no external cause is known for a given fall, there is no reason why one should pretend to know the cause of its sequent rise. The wiser effort would be to seek hypotheses which would account for rise-fall, rise-rise, fall-fall, and fall-rise sequences, one and all, and, until such hypotheses could be subjected to verification, to be content with attributing them to 'accidental' variations.

RHYTHM OF ATTENTION

I have no new data to report bearing on the theory of Voss ['99] that efficiency in adding fluctuates in a rhythm of 2 up to 3 seconds. But since Voss's work has been referred to approvingly by later writers, it may be worth while to show its essential unsoundness.

Voss's study is instructive as a sample of the difficulty of keeping in mind just what a set of measurements means while one is arguing from them. What he does is in essence as follows:

Suppose the times of a series of additions are, in order, in fifths of a second, 3-3-6-5-3-8-3-1-3-1-3-4-3-3-3-3-6-2-3-3-5-4-5-3-3-2-4-2-3-3-4-2-7-3-3.

What Voss calls the duration length of a fluctuation is such a time as, in this case: 3 plus 3 plus 6; or 5 plus 3 plus 8; or 3 plus 1 plus 3 plus 1 plus 3 plus 4; or 3 plus 3 plus 3 plus 3 plus 6; or 2 plus 3 plus 3 plus 5; or 4 plus 5; or 3 plus 3 plus 2 plus 4; or 2 plus 3 plus 3 plus 4; or 2 plus 7. That is, Voss divides the entire work into a series of lengths of time called fluctuations—in this case 12, 16, 15, 18, 13, 9, 12, 12, 9—each ending with an addition slower than any since the end of the last fluctuation. The length of each fluctuation is then always the sum of *two or more* single addition times.

Now, this being the method of scoring and the time of three fifths of a second being the most common, differences of three fifths of a second in the fluctuations *must* be the most common differences. The case is as if one reckoned fluctuations in the income of a man whose salary was almost always raised or lowered by approximately $100.00. The fluctuations would of course almost always vary by steps of approximately $100.00.

Voss, finding such a variation, hails it as important and discusses seriously explanations of this, to him, very remarkable phenomenon, by attention waves and other features of the worker, but finally observes that it is due to the fact that by the artificial conditions of the test and method of scoring one fluctuation must differ from another in length by the length of one, two or more additions.

He clings, however, to the faith that the great frequency of fluctuation lengths from 10 to 13 fifths of a second has some real and unitary cause in the constitution of the worker. He considers the relation to the respiration curve, and to the optimum time between the warning signal and the stimulus in experiments in reaction time, and concludes that the great frequency of fluctuations of 10 to 13 fifths of a second is due to the fact that 'the attention tends to rise to its highest intensity (*Spannung*) in periods of somewhat over 2 seconds,' so that 'the most frequent length of a fluctuation expresses exactly the length of fluctuations of attention found in other investigations!'

There are two flaws in this argument. First, the length of fluctuation which Voss measured is not the length from a point of maximum efficiency to the next maximum nor from a minimum to the next minimum. Such lengths can be found only by somehow defining 'maximum' (say as an addition of ⅖ seconds or under) and 'minimum' (say as a time of over ⅗ seconds), and by them measuring the intervals. Voss's measures are of the intervals between *all* turning-points. Such sequences as 1-2-1, 2-3-2, 6-7-6, or 7-8-7 give end-points for his fluctuations as truly as do the sequences 1-3-5, 4-1-2-6 and 3-1-1-5 in the series 6-1-3-5-4-1-2-6-3-1-1-5.

In the second place, the variation of single addition-times being around such modes as he indicates with such frequencies as he indicates, their appearance in an order *absolutely random* as respects speed, would give substantially as great a preponderance of fluctuations of 10-13 fifths of a second as he found. It seems to have been a misunderstanding of the laws of probability that led Voss to think that such a random order would make fluctuations of 7 fifths of a second the most frequent. It certainly would not.

Since Voss does not give the required data, I cannot calculate exactly what the relative frequencies of different lengths of fluctuation would be by mere chance, but they would certainly show a clustering closely corresponding to that which he explained by the rhythm of attention. I take at random his experiment VIII. Suppose the time for the separate additions to be such that for every addition taking 9 fifths of a second there occurred one taking 8 fifths of a second, two taking 7, three taking 6, six taking 5, eleven taking 4, 73 taking 3, and three taking 2. This is very close to the distribution in Voss's experiment, given on page 409 of his report. Suppose the order of occurrence of these times to be absolutely random. Then about a third of all the 'fluctuations' as defined by Voss would be from 10 to 15 fifths of a second inclusive. This is the fact which he finds (on page 425 of his report) and from which he reasons to a 2-3 second rhythm with the attention-wave as a cause! The same will be found to hold for all the results of the experiments reported by this author.

OTHER RHYTHMICAL FLUCTUATIONS

Swift and Schuyler report a fluctuation of ten or twelve minutes in length in the case of the accuracy of typewriting, but also fail to give data in support of this observation. They say:

> "Examination of the written exercises showed that the errors always came in bunches, and, when in threes, one group was near the beginning, another about the middle and the third toward the end of the half-hour of practice. Sometimes there were only two groups, but after the mistakes were marked with red ink, the alternation of spotted spaces with those free from marks was striking." ['07, p. 310]

Yoakum ['09] has reported as phenomena of fatigue fluctuations in efficiency of about two or three minutes length appearing within the work of a half hour. He says:

"I. There is a decided fluctuation in *errors* and a fluctuation, somewhat noticeable, in the rate of tapping in each experiment. These fluctuations are such as to give 'groups' of errors in any given experiment. These groups reach their maximum at variable periods, but, in general, occur about two or three minutes apart."

"II. The errors when arranged in two classes with reference to the more or less mechanical character they exhibit, follow this grouping in a very peculiar way; to wit, those that are considered more mechanical in their nature, Class II, appear in groups that almost alternate with the first class or 'psychic errors,' and always appear first." ['09, p. 92]

He does not give the details which would enable one to verify or refute these statements. It is not even proved that his subjects showed more rhythm than a chance distribution, plus some special tendency of a slip in efficiency to produce several errors, might give.

WARMING UP

The best definition of 'warming-up' as an objective act is that part of an increase of efficiency during the first 20 minutes (or some other assigned early portion) of a work period, which is abolished by a moderate rest, say of 60 minutes. Such warming-up should show itself clearly in individuals at or near the limit of practice, and, in others, should compound with the effect of practice to make the rise in efficiency especially rapid in the first twenty minutes of work, or the fall (supposing the function to diminish in efficiency) specially slow in this same period. What time is assigned in the definition of warming-up effect is of little consequence to the investigation so long as *some* time is assigned.

In the case of my five subjects in addition it is clear from

FIGS. 10 to 14 that such objective warming-up either completed itself within a minute or so, or was so slight as not to appear consistently and not to appear at all in any considerable magnitude.

There is little or no direct evidence of warming-up in the records got by Oehrn ['95], Lindley ['00], Weygandt ['97], Bolton ['02], Miesemer ['02], or Rivers and Kraepelin ['96]. Possible indirect evidence of it may be got from the finding of Wimms ['07] that 20 minutes of work at simple computation, but in a form involving trying control of the eye's fixations, was more efficient when done in two equal periods with a ten-minute rest between than when either no rest or a 20-minute rest was given. My sixteen subjects ['11] working at mental multiplication of a three-place by a three-place number showed signs of its presence, but not conclusively.

It seems likely, from the cruder observations of daily life, that for many individuals in many functions, there is a warming-up effect as defined, but I am unable, with the data at hand from Kraepelin's pupils or others, to separate out this temporary improvement that comes at the beginning of the exercise of a function after rest, from the more permanent improvement that comes from exercise of the function in general. I am confident that it has commonly been exaggerated. Experiments near the limit of practice are necessary. One must be careful not to mistake the sudden gain from practice in its early stages for warming-up. There is a temptation, for example, to see evidence of the latter in the first day's curves for my five subjects (shown in FIGS. 10 to 14). But the absence of any similar gain in the first half hour in the curves for the last day for these same subjects proves that its presence on the first day was probably a result of practice. Arai, who worked near the limit of efficiency, showed no demonstrable warming-up effect.

It should also be noted that intellectual warming-up in the popular sense refers rather to fore-exercise of *other functions,* in order to get materials and motives with which and by which the given function is to work, than to an intrinsic alteration of it. Such is the case, for example, with the 'poets, artists and composers' of whom Mosso* speaks.

There is also probably often a rapid *relearning,* with consequent rise in the score, during the first few minutes of a practice-period. This is perhaps what is meant by Warming-Up or Incitement by certain writers. It is doubtful whether a rest of sixty minutes or so would abolish this rise in the score resulting from relearning. And it seems more useful to think of a rise due to relearning in the terms which exactly describe it, rather than in the vaguer terms—Anregung, Incitement, or Warming-Up.

ADAPTATION (GEWÖHNUNG)

Adaptation is, so far as I can see, definable objectively only as a slower, longer and somewhat more permanent warming-up effect. I am unable to identify or measure it in the accessible work-curves.

SUMMARY

The essential empirical facts about the curve of mental work seem then to be as follows: Two hours or less of continuous exercise of a function at maximum efficiency, so far as the worker can make it so, produce a temporary negative

* "Seneca has said that one must force one's mind before it will begin to work:

Cogenda mens, ut incipiat,

and Alfieri used to make his servant tie him to his study table. Without going to such extremes, we all know that in any intellectual work whatever we do not get on so well at first as later. In works of the imagination, which demand creative and associative power, this difference is more perceptible than in works of the reasoning faculty or in science, where what is necessary is only the comparison of facts presented to us by nature. Poets, artists, and composers especially feel the need of 'winding themselves up,'" [*Fatigue*, English translation of 1904, p. 299] effect, curable by rest, of not over ten per cent, and in most functions still less than that. Fluctuations of considerable amount occur in any one work period for any one subject, but except for a rise in achievement of approximately four per cent near the end when the date of the end is known, no regularity in them has been proved for any one of them for any one subject in any one sort of work, much less for any one subject in all sorts of work, or for all subjects in any one sort of work. The supposed laws that the very first few minutes and the minutes after a drop in efficiency are periods of specially high efficiency are not supported by the facts. A special gradual increase in efficiency in the first fifteen or twenty minutes is not demonstrable in the case of the simple functions such as addition, mental multiplication, marking words of certain sorts and the like. The fluctuations in a single day's record for a single subject are then in no sense explained by referring them to fervor at starting, fervor after disturbance, fervor after fatigue, incitement or adaptation.

The most important fact about the curve of efficiency of a function under two hours or less continuous maximal exercise is that it is, when freed from daily eccentricities, so near a straight line and so near a horizontal line. The work grows much less satisfying or much more unbearable, but not much less effective. The commonest instinctive response to the intolerability of mental work is to stop it altogether. When, as under the conditions of the experiments, this response is not allowed, habit leads us to continue work at our standard of speed and accuracy. Such falling off from this standard as does occur is, in the writer's opinion, due to an unconscious reduction of the intolerability, by intermitting the work or some parts of it.

SPECULATIVE ANALYSES OF THE CURVE OF WORK

So far in this chapter fatigue, warming-up, initial spurt and the rest have been considered as objective features in a work-curve. Psychologists have also often used these terms as names for mysterious forces which caused efficiency to rise and fall. *Antrieb* is thus an inner fervor or impulse which causes a spurt; *Anregung* is thus an inner incitement which causes the warming-up effect. *Fatigue* is thus the inner cause of diminished efficiency rather than the diminution itself.

A work-curve may then be considered as the gross result of a compounding of all these forces in various degrees, and may be analyzed into imaginary curves, each corresponding to the action of one of these forces. The analysis which anyone makes will depend upon his theories about their separate action, and will display those theories clearly.

Let us then, for contrast with the views which I have suggested in this chapter, examine Kraepelin's analysis of the following work-curve: In six successive 5-minute periods before a 30-minute rest, and in six immediately thereafter, the observed numbers of additions were: 483, 473, 478, 473, 486, 474, rest, 496, 510, 513, 489, 497, 494.

Kraepelin's explanation of the above case ['02] is as follows: The efficiency of the function, apart from practice, fatigue, incitement, adaptation, and fervor from starting, finishing, etc., is 393. The effect of practice is to raise this by 56, 40, 37, 34, and 32 from one to another of the first six periods respectively; of this total practice gain of 199, 149 is lost during the 30 minutes of rest;* during the last half hour the practice gains from one to another of successive 5-minute periods are 31, 29, 27, 26, 25. Thus work under the influence of only practice and the 30-minute rest would, in Kraepelin's opinion, give a curve of 393, 449, 489, 526, 560, 592 (rest here), 443, 474, 503, 530, 556, 581 as shown in the upper continuous line of FIG. 17.

The effect of fatigue is to lower efficiency from period to period successively by 41, 41, 41, 41, 41, a total loss of 205; the rest raises it by 197 (in divisions per 5 minutes of 80,

* Or strictly from the last 5-minute period before to the first one after the rest.

45, 35, 26, 7, 2, 2); after rest fatigue again lowers it in successive steps by 41, 41, 40, 39, 38. The work under the influence of only fatigue and the rest would, in his opinion, give a curve of 393, 352, 311, 270, 229, 188 (rest here), 385, 344, 303, 263, 224, 186, as shown in the dash line of FIG. 17. Practice, fatigue and the rest together would give the middle curve of FIG. 17.

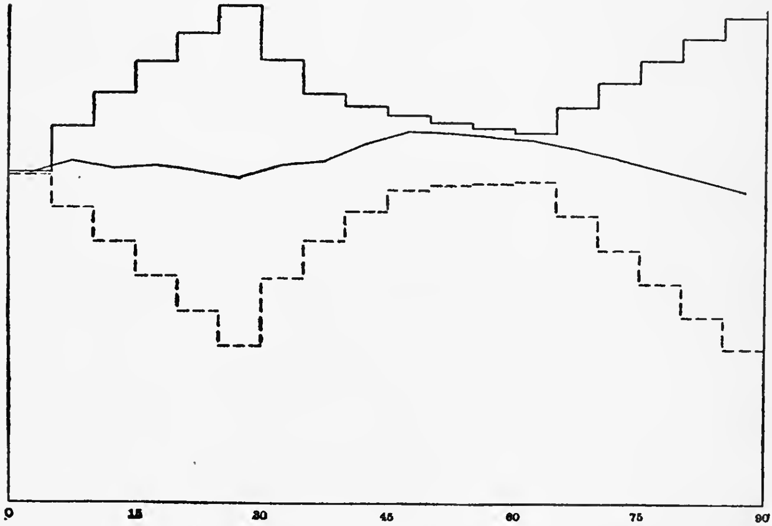


FIG. 17. Kraepelin's Estimates of the Effects of Practice and Fatigue during 30 Minutes Continuous Adding, 30 Minutes Rest, and 30 Minutes Continuous Adding. The horizontal scale is of time; the vertical scale is of number of additions. The upper line gives the efficiency of the function by practice; the lower line gives the efficiency of the function by fatigue; the middle line gives the combined result of both.

Incitement or 'warming-up' is supposed by Kraepelin to raise efficiency from period to period by 35, 5, 5, 0, 0, all the 45 being lost during rest and the gains after rest being 35, 10, 0, 0, 0. The effect of adaptation is supposed to be, from period to period, 30, 5, 5, 5, 5; 5 of this total 50 is lost during rest; after rest the gains due to it are 5, 4, —8, 10, 1. This effect of incitement is shown in FIG. 18 by the line of long dashes; that of adaptation by the line of short dashes.

The effect of fervor at starting is to make the first five-minute period better by 90, and the first one after rest better by 16, than they would otherwise have been.

Minor fluctuations make the third period 4 worse, the fourth, 15 worse, and the sixth, 8 worse, than they would otherwise have been. The ninth and tenth periods are so influenced to the extent of —1 and —2. Fervor at the approach of the end makes the last period of the whole series

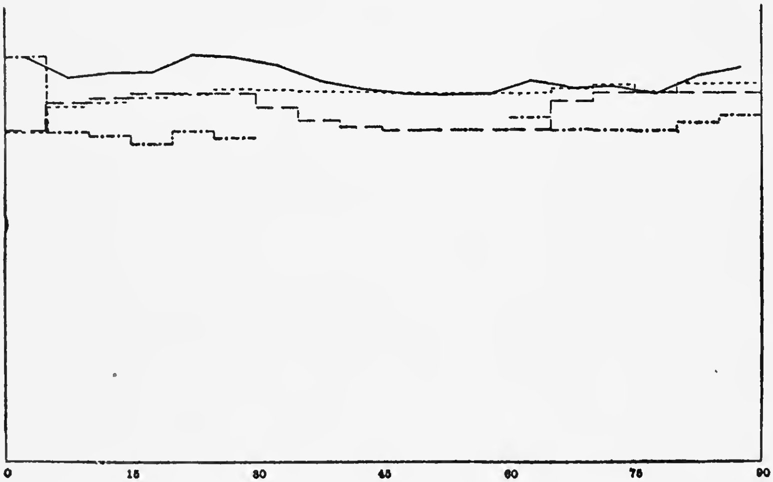


Fig. 18. Kraepelin's Estimates of the Effects of Incitement, Adaptation. Fervor at Starting and Fervor at the Approach of the End during 30 Minutes Continuous Adding, 30 Minutes Rest, and 30 Minutes Continuous Adding. Scaling as in Fig. 17. The long dash line is for Incitement; the short dash line is for Adaptation; the dash and dot line is for Fervor at starting, ending, disturbance, etc.

better by 18, and the next to the last better by 9, than they would otherwise have been. The effect of these various *Antriebe* is shown by the dot-and-dash line of Fig. 18.

Incitement, adaptation, beginning and end fervor, and minor fluctuations all together thus would with the rest give a curve of 483, 448, 464, 463, 483, 480 (rest here), 454, 478, 491, 482, 503, 513, as shown by the continuous line of Fig. 18.

This analysis by Kraepelin is an instructive illustration of the danger of speculative *ex post facto* explanation, for it is certainly highly improbable in almost every one of its main features. If a man of Kraepelin's gifts and training could so err, others surely can be rarely trusted to imagine the causes of a given curve of work.

The first error in the analysis is the assumption that thirty minutes' continuous exercise will add 50 per cent to efficiency in addition in an educated adult who already adds any one-place to any two-place number in three quarters of a second. No expert in the psychology of practice and of reaction-time would dare to assume so great a gain in so short practice in a function already so far improved.

The second error is the assumption that, in an interval of thirty minutes rest, three-fourths of the practice gain is lost. All studies of practice in the case of functions like addition show great permanence of effect after many hours and even days.

The only explanation of their results that would be consistent with any such loss within thirty minutes as Kraepelin here assumes, would be that, even within five minutes or less, there is an enormous decrease in efficiency due to fatigue, so that the net apparent gain from day to day was only a small fraction of what is gained in one day, but lost by fatigue—this in experiments in which the best efforts of the investigator were directed toward avoiding the possibility of any fatigue at all!

A third error is in the assumption that the total practice effect of an hour's addition broken by a thirty-minute rest is inferior to that of half an hour's addition alone (+ 188 and + 199). It might be that in a few individuals no practice effect would come from a second half hour of addition after such rest, but it is enormously unlikely.

A fourth error is the assumption that the fervor of beginning has five times* the potency of the fervor from the approaching end of work. There is, on the contrary, a pre-

* Three times from one point of view.

ponderance of evidence to support the belief that, of the two, finishing spurt is more influential than initial spurt in mental work generally and in addition in particular.

A fifth error is the assumption that the fatigue from thirty minutes of work will, apart from practice, incitement and adaptation, reduce efficiency by over half. We can reduce the influence of practice, incitement and adaptation by testing individuals at or near the limit of practice, and in the second half of a ninety-minute period. We do not then find a reduction at all comparable to the assumed 52 or 53 per cent.

A sixth error is the assumption that a half hour's rest will cause seven and a half times as great a percentile loss of the effect of practice as it does of the effect of adaptation! Adaptation, to mean anything distinguishable from practice, should mean an improvement that is *shorter*-lived, that can occur again and again day after day as practice cannot, because it dies out during lack of exercise of the function as practice does not.

The simple fact is that the supposed fatigue of fifty per cent in thirty minutes and the supposed initial spurt of twenty-five per cent in two or three minutes are both myths. The enormous practice by a half hour's adding, and the enormous loss of its effects in a half hour's rest, are myths unconsciously invented to support the first myth; the warming-up and the adaptation are modelled into the mythical form necessary to support the second myth.

With increased knowledge and greater watchfulness against inconsistent and improbable suppositions, a psychologist today could improve upon Kraepelin's analysis, but any and all such explanations after the fact are exceedingly risky. In the present state of knowledge it is far better to analyze a work curve by experiment than by deduction. If one thinks that a knowledge of the approaching end produces five per cent gain in the last ten minutes of an hour, he should have the subject work at the task both with and without knowledge of the approaching end and with all other conditions identical.

The difference then found is the effect of knowledge of the approaching end. If certain drops in the curve are thought to be due to disturbances, the disturbances in question should be withheld and applied or increased and decreased experimentally until what difference they actually do make has been measured. The influence of initial spurt, warming up, adaptation, and fatigue upon the early course of the work-curve should be made the subject of speculation only after each of these mythological entities has been defined as some state of affairs in or out of the organism that can be observed to happen or not to happen, or to happen in varying measurable amounts, and after records have been made of the subject's achievement as related to different amounts and combinations of these states of affairs. At present it is purely fantastic to replace the given fact that in the first twenty minutes such and such a product is achieved, by (*a*) estimates of what that product would have been apart from initial spurt, that is, what work would have been done in the beginning if there had been no beginning, (*b*) estimates of what that product would have been apart from 'warming-up' (that is, of what work would have been done at a new task if the task had not been new), and (*c*) estimates of what that product would have been apart from fatigue (that is, of what work would have been done if no work had been done)!

THE CURVE OF SATISFYINGNESS

All our measurements so far have been of the quantity and quality of the product, not of the satisfyingness of the process. Of the latter, indeed, there have been only occasional and very crude reports. No one has ever made the experiment of arranging with workers that they should work at least two hours, and be given, say, two cents a minute for every minute that they continue at maximal exertion beyond two hours, being compelled to pay two cents for every minute less than two hours if they stop work. Nor has any experimenter equated the requirement of *X* production or the privilege of *Y* rest at given stages in a period of production with any measure of value whatever.

In ordinary life, such equating is of enormous importance. Having, say, a thousand sums to compute, one does them all rapidly and continuously, resting at the end; or at moderate speed, not resting at all; or rapidly at first, and then more and more slowly; or one inserts a thousand rests of two seconds; or eight or ten rests of three or four minutes; and so on through the infinite variety of ways of administering one's mental production. What one does is determined in large measure precisely by balancing the satisfyingness of money-rewards, free time, following familiar customs, and the like against the annoyingness of this or that feature of the process. John, who is annoyed by hurry, inserts his rests bit by bit. James, who is more annoyed by the lack of free time for some cherished pursuit, saves his rests till the product is complete.

Anyone who undergoes experiments in working mentally as continuously and perfectly as he can for a long period can get some rough idea of the curve of satisfyingness for his own case, with the function in question, in the circumstances of competing desires in question. In the general accounts of mental work and fatigue, the impressions obtained thus, or from the ordinary experiences of life, play a part. The notion of Warming-Up, for example, has included the diminishing annoyance of a process due to the progressive inattention to competing desires, interest in achievement, and the like. So, also, the effect of knowledge that the end is nigh in producing a spurt is attributed, and probably rightly, in part to the fact that the satisfyingness of reaching the end, and of using one's last chance to do one's utmost spreads to make the process itself more satisfying.

The curve of satisfyingness need not follow the curve of achievement. The slight amount of the decrease in efficiency, due to continued exercise of a function under the condition that one shall do his utmost throughout, may be accompanied by a very great decrease in the satisfyingness of the process. The very individual who after five or six hours, adds or multiplies more rapidly and accurately than ever, may be in a condition which would in ordinary life make him stop the work on grounds of absolute unfitness to continue. The supposed unfitness would not be an inevitable inefficiency in the function itself, but the effort, tension and misery for which its unsatisfyingness was responsible. Indeed, the less a man was fatigued in the sense of becoming unable to produce, the more he might be fatigued in the sense of finding the process intolerable.

Although little can be said about the effects of any given continued task on satisfyingness, such are of very great importance. In ordinary life the amount, rate and changes in rate of increase or decrease in the efficiency of a given function are not determined in any simple mechanical way by the amount of energy possessed at the start, the opportunity for it to be spent, and the lengths of rest periods devoted to its recuperation. Nor are they determined by mysterious tendencies—*Anregung, Gewöhnung, Arbeitsbereitschaft, Ermüdungs-antriebe, Störungs-antriebe,* and the like. They are determined, as are any responses of the animal, by its original tendencies, past experiences and present attitude, including the tendencies to be satisfied and to be annoyed by this and that state of affairs.

A man does not, by beginning to add, open a valve which releases mental energy at a rate depending on the store of it possessed. Inactivity does not necessarily restore the energy. Nor does the valve work wider and wider open by *Anregung,* or widen and contract every few seconds by the *Rhythm of Attention,* or open very wide by a strange foresight just before it is to be closed. A man's behavior during two hours of adding is a series of responses to whatever of the original situation persists plus the new elements due to each stage of the work. These responses are conditioned only slightly by such changes in the animal as can properly be likened to a lessening of a fund of energy. The appeals of ungratified impulses as they weaken by inattention or grow stronger by the lapse of time, the loss of the zest of novelty as the same process is repeated, the sensory pains from strained posture, misuse of the eyes and the like—are all as truly effective in determining the work-curves of ordinary life as is the mere amount of time employed or product produced. But these act primarily on the satisfyingness of the process and only indirectly on the quality and quantity of the product.

CHAPTER IV

The Influence of Continuous Mental Work, Special or General, Upon General Ability

It is really idle to inquire whether fatigue is specific or general—whether, that is, continued work at one function diminishes efficiency in only it, or in all functions equally. We do not have to choose between these alternatives. The first is almost always, and the latter always, false. It is a separate problem to tell for any given loss in efficiency of any function due to its exercise without rest just what the effect upon every other function will be. Some functions will suffer little or not at all; others, much. The real questions are: "*How much* does continued work at any one or any combination of tasks diminish efficiency for any other task?" and "How does it?"

The same doctrine of transfer by identical elements noted in the case of the influence of improvement in one mental function upon the efficiency of other functions is applicable here to the influence of diminished efficiency. However, as was noted in the case of practice, we usually lack the knowledge by which to know beforehand in what respects and to what extent two functions are physiologically identical. Moreover, we lack, in the majority of cases, knowledge of how the various elements of a function share in the total loss in efficiency. Consequently, if a function, say, adding, loses one fifth in efficiency as a result of five hours of work, it is as yet impossible to prophesy the resulting loss in another ability, say, to memorize nonsense syllables, save very roughly.

The clearest cases of identical elements important in the transfer of fatigue, are the headaches and deprivations. A headache produced by five hours of mental multiplication may act equally to diminish efficiency in writing poetry. The deprivations from sleep, exercise, sociability, games and the like are common elements of very many different forms of mental work. Just as learning not to be distracted by certain impulses is a means of improvement common to many functions, so the increased urgency of these impulses due to long deprivation may be a means of decreasing efficiency in many functions. In many of the actual tasks of school, professional, business and industrial work, eye-strain is an important identical element. Factors essentially irrelevant to fatigue itself—notably, excitement, worry, sleeplessness and loss of appetite—may appear in connection with work at one task and diminish efficiency for many, or even all, other tasks.

Students of mental work and fatigue have not, however, commonly thought of the problem as a series of special problems of the influence of a given amount of work with a given function or functions upon the temporary efficiency of each of countless others. They have thought of mental work vaguely as the work of 'the mind' or 'the brain,' and have openly or tacitly accepted one or another form of the theory that continuous work reduced some supply of mental energy. They have commonly assumed therefore that any work must reduce the efficiency of the mind for all work.

They seem to have expected also that any work would reduce the efficiency of the mind *equally* for all work. The questions in their minds have been, "Does this or that work temporarily enfeeble the mind?" and "Does such and such a test measure power to work?"

There resulted about a dozen investigations each attempting to measure the effect of a more or less well defined amount of mental work upon the efficiency of some convenient sample of the mind's operations. But the investigators, obsessed by the too simple theory of a supply of mental energy, usually neglected to eliminate the influence of practice and interest and to verify their right to attribute certain losses in efficiency to the amount of preceding continuous work. Consequently it is difficult to tell exactly what their results really signify without making elaborate accessory experiments.

With intelligent allowance for the effect of practice, interest, conventional habits, suggestion, and the like, however, the main facts of importance concerning 'general fatigue,' especially in school children, can be demonstrated. I shall first present samples of the experimental results that have been obtained and then summarize the conclusions to which they lead.

EXPERIMENTAL RESULTS

Samples of the facts are, briefly, as follows:

Sikorski ['79], testing the same children first before school and then after school in writing from dictation, found the average percentage of wrong letters to be for six grades:*

Before school	1.24	1.21	.72	.66	.61	.46
After school	1.57	1.45	1.03	.94	.81	.80

The central tendency is thus toward a third more errors in the late tests (.7 and .10).

I trust that the reader is not so unsophisticated as to assume that the above figures, even if taken at their face value, show an efficiency before school 1.33 times that existing after school. They as truly mean that, since about 99.3 per cent of the letters were correct in the morning, and about 99.0 per cent after school, the efficiency before school was 1.0033 times that existing after school.

Bolton ['92] reports the results of tests, arranged in coöperation with Dr. Franz Boas, to measure the number of digits that could be remembered after a single hearing (using lists of 4, 5, 6, 7 and 8) early and late in the school day. He tested 136 pupils four times in the morning, and 219 pupils four times—first, late in the session; second, early

* Sikorski carelessly took no account of the speed of work, so that he did not measure the efficiency of the function at all.

the next morning; third, late in that day's session; and fourth, early the next morning. Each test comprised twelve series of figures. The 219 pupils did almost exactly as well at the end as at the beginning of the session, though the combined result of practice and novelty should have made the second and fourth tests better than the first and third. The 219 pupils also improved as much from morning to night as from night to morning, and as much as the 136 pupils used as a check improved from one test to the next. Indeed, the data show a slight apparent advantage in the late period.

Laser ['94] tested four classes of pupils in school with addition and multiplication, of the type described on page 21, at the beginning of each hour of a school day. A time limit of ten minutes was set. The results were as shown in Table 13. On the whole the work improved with the course of the day's work. If his scores are reduced to amount achieved, accuracy being equal, by allowing five operations for each error, the results for the five periods are in the following proportions:

	Class A	Class B	Class C	Class D	Average of all Classes
1st hour	100	100	100	100	100
2nd hour	119	101	114	114	112
3rd hour	123	101	117	120	115
4th hour	119	107	123	118	117
5th hour	122	124	120	123	122

Bergström ['94] found, in a test of twelve days, that his efficiency in reading numbers, adding and multiplying, suffered no diminution in efficiency as a result of four hours' maximum work at translating a foreign language.

Bourdon ['95, p. 160 ff.] found in marking *a*'s in pages of mixed small letters the following results: In the first test (4 min.), 1693 letters were examined, 216 *a*'s were marked, 7 *a*'s were omitted. In a test (4 min.) after 12 minutes intervening work in marking *a* and other letters (*a* and *e*, 4 min.; *a, e, l, t,* 4 min.; *a, e, l, t, o, k,* 4 min.), 1817 letters were examined,

TABLE 13.

EFFICIENCY OF SCHOOL PUPILS IN EASY COMPUTATION AT EACH HOUR OF THE SCHOOL SESSION. After Laser, '94.

Sex and average age of the class	Period of the School session	Average number of additions or multiplications	Average number of errors
G. 11y. 9m.	1	207	7.7
	2	261	10.0
	3	275	11.0
	4	271	12.1
	5	269	10.2
G. 11y. 1m.	1	188	5.0
	2	195	6.3
	3	198	6.4
	4	211	7.0
	5	241	6.4
B. 11y. 6m.	1	126	4.5
	2	148	5.0
	3	156	5.8
	4	165	6.1
	5	164	6.7
B. 10y. 7m.	1	99	3.3
	2	123	4.9
	3	143	7.5
	4	139	7.1
	5	143	6.6

238 *a*'s were marked, 3 *a*'s were omitted. A week later, with the same experiment, save that the type was smaller, the results were: In the first 4 min., 1591 letters were examined, 208 *a*'s were marked, 3 *a*'s were omitted. In 4 min., after 12 min. other marking, 1799 letters were examined, 233 *a*'s were marked, 4 *a*'s were omitted.

Friedrich ['97] tested a class of 51 pupils of an average age of 10 years, on eleven occasions during a period of six weeks, using dictation, addition and multiplication. The last two were of the type $\frac{275831406}{69413258}$ and $\frac{27583140}{2}$ (or 3. 4. 5. or 6).

With the dictations there was an increase in the number of errors in later over earlier periods, and in the same periods without previous rest over earlier periods with rest; but as no record was kept of the speed of the work, efficiency cannot be measured. From the results of other similar tests, it is probable that the speed increased. The results as to accuracy are given in Table 14.

Friedrich gave twenty minutes time for about 206 single additions or multiplications. Consequently all but the very slowest pupils finished all the examples before the time was up, so that we have no record at all of the speed of the work after the first test, and only a very imperfect and distorted record for that. Now it is known from the results of other workers that in repeated tests with such simple additions and multiplications, a child tends to work more and more rapidly at the cost of accuracy. The number of errors is then certainly a wrong measure of efficiency. Apparently, if he had recorded the time taken, he would have found, even allowing a discount of so many as ten additions for each error, no decrease whatever in the efficiency-scores for the tests late in the school day. His gross results as to accuracy are given in Table 14.

TABLE 14.

THE RESULTS OBTAINED BY FRIEDRICH CONCERNING THE ACCURACY OF SCHOOL WORK AT DIFFERENT PERIODS OF THE DAY.

	Time of test	Letters, etc., written in Dictations		Figures of sums and products in Computations	
		Per-cent right	Per-cent wrong	Per-cent right	Per-cent wrong
Morning Session	Before 1st hour	99.8	.2	98.9	1.1
	After 1st hour	99.6	.4	98.4	1.6
	After 2nd hour and 8 min. rest	99.3	.7	98.0	2.0
	After 2nd hour	99.2	.8	98.0	2.0
	After 3rd hour and two 15 min. rests	99.4	.6	98.1	1.9
	After 3rd hour and 15 min. rest	99.0	1.0	97.8	2.2
	After 3rd hour	99.0	1.0	97.7	2.3
Afternoon Session	Before 1st hour	99.8	.2	98.1	1.9
	After 1st hour	99.2	.8	97.9	2.1
	After 2nd hour and 15 min. rest	99.4	.6	97.9	2.1
	After 2nd hour	98.9	1.1	97.6	2.4

As with Sikorski's results, so with Friedrich's, the number of errors has been carelessly taken as a direct measure of inefficiency. Binet and Henri, for example, print a diagram like FIG. 19,* which gives the impression that efficiency decreases enormously with the progress of the school-day and the absence of rests. Even supposing the speed of work to have remained constant, this diagram is very misleading.

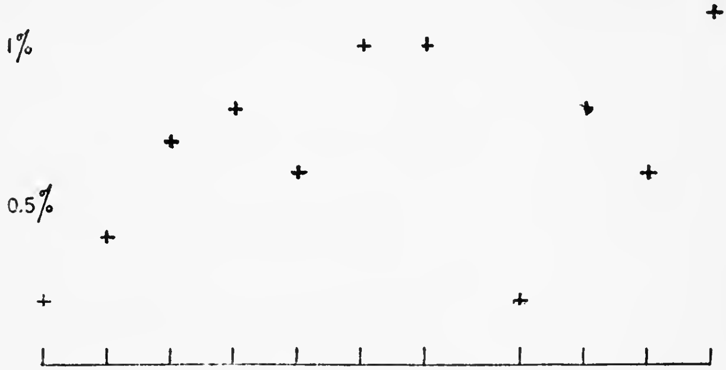


FIG. 19. Friedrich's Results, Scored by Percentages Wrong.

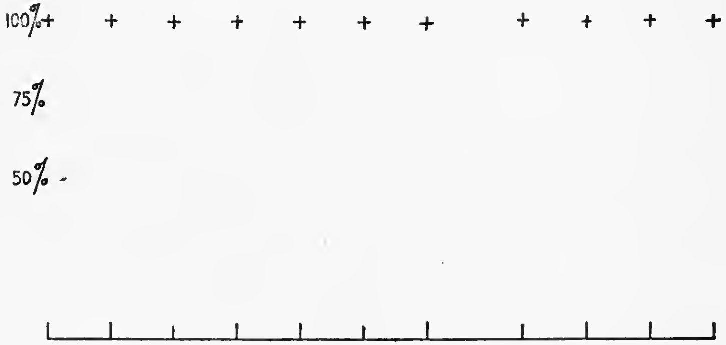


FIG. 20. Friedrich's Results, Scored by Percentages Right.

Efficiency is as properly measured by the number written correctly as by the inverse of the number written incorrectly. If the former measure is used, the diagram becomes FIG. 20. FIG. 19 and FIG. 20 measure *absolutely the same fact*. The

* Binet's diagram is not identical with FIG. 19, but is even more misleading, since he adds to the errors in the finished work, the errors made *but corrected by the pupils during the dictation*. An error noted and corrected by the pupil as he works can hardly be ranked equivalent, as a sign of inefficiency, to one left in the result.

case is as if we had a man's gains and losses in trade for consecutive hours of the day. Suppose him to have made 998 dollars and lost 2 dollars the first hour, to have made 996 dollars and lost 4 dollars the second hour, and so on. Both gains and losses should figure in estimating his efficiency. Now in the case in hand, we do not know how much more efficiency is required to copy correctly 998 out of a thousand letters than to copy 996. The first performance is, however, surely not two times as efficient.

Since Friedrich does not state the order in which his tests were made, we cannot tell what part practice and loss of interest played in these results. The influence of either, in a test of a regular feature of school work by the regular teacher, with no information given to the pupil of the results, was probably very slight.

Under directions issued by Ebbinghaus ['97] the teachers of certain classes in Breslau tested their pupils at the beginning of the school day and at the end of the first, second, third, fourth and fifth hours. On one day the six tests were with immediate memory of series of digits (2 each of 6, 7, 8, 9 and 10 digits); 2 weeks later there were six tests each of ten minutes length, with such examples as $\begin{array}{r}46753294\\+62857923\\\hline\end{array}$ and $\begin{array}{r}57835694\\\times\quad 2\\\hline\end{array}$ (or x 3 or x 4 or x 5 or x 6). Two weeks later there were six tests, each of five minutes length, in filling in omitted words and syllables in selected passages. The pupils ranged from about 9 to about 19 years in age. The results are subject to the influence of practice and interest and to that of the disturbing factors (notably conventional habits and personal suggestions) due to the administration of the tests by teachers unskilled in experimentation.

In memory of digits, whatever decreased efficiency the work of the school-day and the loss of interest due to repeating the test may have produced was over-balanced by the influence of practice and other factors. The number of digits correctly written was about one tenth greater in the last three than in the first three tests. The facts in detail are given in Table 15.

In addition and multiplication there was no demonstrable difference in efficiency after the first test. The achievement in that was about one fourth *less* than that in any of the others. If we subtract five single additions or multiplications

TABLE 15.

EFFICIENCY IN MEMORIZING DIGITS. EARLY AND LATE IN THE SCHOOL SESSION. After Ebbinghaus, '97.

Class		At the beginning and end of the first hour. Average number of figures correctly given out of a possible:			At the end of the fourth and end of the fifth hours. Average number of figures correctly given out of a possible:			Per cent which late achievement is of early achievement.
		32	36	40	32	36	40	
Unterprima	A	29.7	33.0	31.9	31.0	34.2	33.2	104
Untersekunda I	B	30.2	32.5	30.7	30.8	34.1	35.8	108
Obertertia I	C	29.0	29.5	29 8	30.1	32.8	35.1	111
Quarta I	D	29.5	30.8	32.5	28.3	31.1	31.6	98
Sexta I	E	23.6	23.7	22.4	27.9	28.6	28.3	122
I Girls	F	28.2	29 7	27.9	29.9	33.0	34.4	113
V (a) Girls	G	25.5	26.3	26.0	24.2	24.1	23.6	92

for each error, the central tendency is for the efficiencies at the different periods to be as 72, 96, 103, 103, 97, and 101. The detailed results are given in Table 16.

In the tests in filling omissions in passages the work toward the end of the school day appears less efficient, that of the last three tests being on the average from 5 to 20 per cent below that of the first three tests, according to how 'amount done' and 'errors made' are equated. The detailed results are given in Table 17. Ebbinghaus is unwilling to use these results as a measure of fatigue, because of the possible inequality of the passages in difficulty and the very great differences in the results amongst different classes.

TABLE 16.

EFFICIENCY IN COMPUTATION IN THE COURSE OF THE FORENOON SESSION. After Ebbinghaus, '97, p. 442. (Number done, number of single additions or multiplications.)

	Beginning of Session		End of 1st hour		End of 2nd hour		End of 3rd hour		End of 4th hour		End of 5th hour	
Boys	Number done	Percent of errors	Number done	Percent of errors	Number done	Percent of errors	Number done	Percent of errors	Number done	Percent of errors	Number done	Percent of errors
Obersekunda	248	1.1	316	1.3	273	1.1	327	1.8	239	1.6	336	2.3
Untersekunda I..	224	0.8	318	1.1	342	1.0	367	1.5	335	1.4	323	1.2
Obertertia I.....	213	1.3	281	1.6	312	1.7	318	1.9	289	2.2	305	2.0
Untertertia I	181		258		291		286		271		268	
Quarta I..........	190	1.5	238	1.8	277	2.0	280	2.0	242	2.3	272	2.0
Quinta I..........	183	0.8	239	1.7	288	2.0	284	2.0	262	1.5		
Sexta I..........	168		221		239		184		219		234	
Girls												
I	241	0.9	366	1.6	353	1.5	356	1.7	344	2.2		
IVa..............	133		190		217		231		229		255	
Va...............	96		144		176		173		173		175	
VIa..............	136		155		181		182		187		163	
Average..	183	1.1	248	1.5	268	1.6	272	1.8	254	1.9	(259	1.9)

All three of his experiments are subject to disturbing factors, due to the administration of the tests by teachers unskilled in psychological methods, and to the varying attitude of the pupils, and so do not allow any safe conclusions.

TABLE 17.

EFFICIENCY IN FILLING MISSING WORDS AND SYLLABLES IN A PASSAGE IN THE COURSE OF THE FORENOON SESSION. After Ebbinghaus, '97, p. 445.

			Beginning of session		End of 1st hour		End of 2nd hour		End of 3rd hour		End of 4th hour		End of 5th hour	
	Number of pupils	Average age	Number of syllables	Percent of errors	Number of syllables	Percent of errors	Number of syllables	Percent of errors	Number of syllables	Percent of errors	Number of syllables	Percent of errors	Number of syllables	Percent of errors
Boys														
Unterprima	24	18.0	63	11	67	10	54	13	67	13	68	18	70	18
Obersekunda	26	17.8	71	9	68	11	54	13	82	13	..	..	..	..
Untersekunda 1	23	17.1	64	14	66	15	57	19	59	19	49	21	60	28
Obertertia 1	37	15.5	48	21	55	18	43	24	57	20	48	26	64	23
Obertertia 2	21	15.3	47	21	52	23	40	30	44	29	46	28	73	23
Untertertia 1	24	14.4	70	5	62	8	66	12	67	12	70	10	67	13
Untertertia 2	18	14.0	55	8	45	16	58	14	56	17	67	14	70	11
Quarta 1	35	13.2	48	9	41	11	53	17	46	23	46	24	59	20
Quarta 2	27	12.4	41	18	38	22	39	32	53	26	50	34	51	20
Quinta 1	27	12.1	47	16	42	22	51	30	46	28	41	35	..	..
Quinta 2	16	11.5	48	18	33	28	47	32	34	33	37	38	..	..
Sexta 1	25	10.7	43	15	28	37	32	42	36	28	27	34	28	39
Girls														
I	36	15.6	52	20	53	22	43	22	..	..	50	32	..	..
II a	24	14.4	39	19	45	30	39	27	34	30	44	28	..	..
IV a	27	12.9	36	10	36	16	27	30	36	25	35	16	40	19
IV b	31	12.6	47	8	44	17	37	27	48	21	43	26	48	21
V a	26	11.7	35	12	37	19	30	27	26	30	30	25	41	26
V b	25	11.5	34	12	37	14	30	33	28	31	28	24	34	21
VI a	33	10.6	20	32	20	40	15	59	13	56	17	58	16	61
VI b	30	11.0	26	27	23	32	17	54	23	52	18	48	..	..

Ritter ['00] tested the ability of pupils in secondary schools to mark letters and words of certain sorts (a double task being imposed each time, such as to mark every s and also every preposition), and to take sensible sentences, lists of unrelated words, and lists of digits from dictation. The tests were given at various times in the school session. In the case of the cancellation tests he gives no measures of amount. In the case of connected dictation he found no change after work. In the case of the unrelated words, there was on the whole a slight falling off. In the case of the digits there was not.

The author ['00], eliminating the influence of both practice and novelty by the simple expedient of never giving the same test twice to any individual, found the efficiency of school children (in adding,* multiplying,† marking misspelled words on a page of print,‡ memorizing lists of 10 digits, 5 nonsense syllables,§ 10 letters and 6 simple forms,¶ and counting dots) substantially the same at near the end as at the beginning of the school session.

The children who were tested with one function early in the school session were tested with another function late in the session. Thus no child ever repeats any test, and all influence of practice and novelty is avoided. At the same time any influence upon the results from the accidental superiority of one half of the children to the other half can be detected and allowed for. The number of children taking a test varied from 240 to 700. Since all the tests were given

* With examples, each of five four-place numbers.

† With 9 examples like $\frac{7986}{4523}$

‡ With a passage like:

> After waiting some time Captain B— and myself walked acros the rice fields to the shad of a tree. There we herd the trumpett of an elephant: we reshed etc.

§ With the following:

> *ba ni su et ko* and *ig fa tu le ro*

¶ The forms being:

by the author and his assistant, each one's time being equally divided between 'early' and 'late' tests, the only factor left to produce any difference is the difference of time of day, with whatever differences in amount of school work and other factors it implied.

The following facts appeared:

Experiment 1.—Those who did the multiplication work late in the session did 99.3 per cent as much and made 3.9 per cent more mistakes than those who did it early, and numbered 64 who misunderstood or grossly failed in the test against 56 among the early ones.

Experiment 2.—The test in marking misspelled words which was given early to those who did the multiplication late, and *vice versa,* shows the following results for those who took it late as compared with those who took it early. Relative amount of page covered, 99.0 per cent; relative number of words marked, 105.0 per cent; relative number of words marked improperly, 97.9 per cent. Thus the decrease in ability shown in test 1 is offset by an equal increase in the case of test 2.

Experiments 3 and 4.—Tests in memorizing figures were given to four classes early and four late. Taking both together, we find that the pupils who did the work late in the forenoon or late in the afternoon memorized almost 2 per cent more than those who had the work early. But with tests in memorizing letters and nonsense syllables, where the pupils were reversed, we find the late pupils memorized only 98 per cent as much as the early with test 4, and 99.8 per cent as much with test 7. So here again the balance is practically equal.

Experiments 5 and 6.—The test in memorizing forms was given late to only half the total group of pupils; and the set who took it early were shown by the other tests to be a little more intelligent. Here the late pupils did only 94.6 per cent as well. That this was wholly due to a difference in average ability is also witnessed by the fact that when 75 per cent of these scholars were tested in counting dots (those who memorized forms early counted dots late, and *vice versa*) the scholars taking the test *late* did much more than five per cent better.

The author ['00] also carried out series of tests with addition and mental multiplication of two-place and three-place numbers, with several subjects. The tests, involving from 2 to 12 minutes time, were given before work and after a full day's work of study or teaching, under such conditions as to equalize practice effect. There was a great decrease in inclination toward the mental multiplication of a three-place by a three-place number in the evening tests, but no appreciable decrease in efficiency at it. Adding long columns of figures required about a sixth longer, probably because of eye-strain. The summed results were:

Mental Multiplication.

Subject T.	Before work, 4630 seconds, 3 errors. After work, 4505 seconds, 9 errors. The time required (with equality of errors) after work was 102 per cent of that before work.
Subject E. T.	Before work, 3019 seconds, 29 errors. After work, 2823 seconds, 21 errors. The time required (with equality of errors) after work was 90 per cent of that before work.
Subject M. T.	Before work, 2805 seconds, 6 errors. After work, 2689 seconds, 17 errors. The time required (with equality of errors) after work was 102 per cent of that before work.

Addition.

Subject K.	Before work, 2735 seconds, 18 errors. After work, 3280 seconds, 21 errors
Subject T.	Before work 659 seconds, 7 errors. After work, 754 seconds, 8 errors.

Using the author's method of eliminating the effects of practice and novelty, King* measured the efficiency of children at approximately 9.30, 10.30, 11.30 A. M., and 1.30 and 2.30 P. M., the school sessions being 9-12 A. M. and 1-3 P. M. Five hundred and ninety fifth-grade children were measured (about 120 at each period) with respect to adding five two-place numbers, multiplying four-place numbers by four-place numbers, and answering the questions of the blank reproduced below. The children worked in classes under the supervision of the regular teacher and Miss King, the test being an unanticipated interruption of an ordinary school day. Three classes were measured at each period.

Answer these Questions as well as you can.
Do not hurry, but do your best work.

1. Write after each of the following the name of the plant or animal which produces it:
   Indian Meal............Mutton............Ham............
   Flour..................Linen.............Wool............
   Bacon..................Sugar.............
2. Name three ways in which the Hudson River is valuable.
3. Name four things you could not have if there were no cows.
4. Is it good to have public parks in large cities? Give two reasons for your answer.
5. The following reasons have been given to prove "Which is more useful, iron or gold?"
   a. Gold is made into jewelry.
   b. Rings of gold cost more than iron rings.
   c. Iron is used to make stoves, frames of buildings, machinery, engines, cars, horseshoes, nails, and many other things.
   d. Some money is made of gold.
   e. Iron is black.
   f. Gold is pretty.
   Mark with a (X) cross the reasons you think are good.
6. Name two outdoor games you could play in January in Florida: Name two you could play in Canada.

* In a study as yet unpublished. Schuyten ['04] also has adopted means to distinguish the effects of the intervening work from the effects of the second trial with the test. He gave the second trial with the test to some pupils before the work-period and to others after it, and verified the author's hypothesis that the degree of interest in a test is one main factor in the efficiency shown by children in it.

The addition tested the speed and surety of a long-formed habit; the multiplication, the power of avoiding errors in a complex and unfamiliar operation (children of the 5th grade having had little or no experience in multiplying with higher than three-place numbers): the other experiment tested a mixture of knowledge and power to use knowledge.

The results, together with a convenient presentation of the relative efficiency of each group of children, are presented in Table 18. There are clearly no marked differences in ability to do mental work. The morning session is, perhaps, a little more favorable, especially for the less interesting work. The work of the morning session produces no deficiency by the middle of its last hour. The last half of the afternoon session seems a little more favorable than its first half. The facts throughout verify the author's conclusion that the mental work of an elementary-school day produces no deficiency in ability to do mental work.

Winch ['10] has found that pupils in evening schools who have already completed a full day's work in shop or office, do progressively less well in the course of the evening-school session. His measurements included efficiency in solving problems, computation, memorizing the gist of passages and writing shorthand. On the whole, the achievement of pupils tested late in the evening session was a sixth less than that of pupils of equal ability who were tested early in the evening.

Winch ['11] also found, with four classes of children, that arithmetical problems were not solved so well when a group worked at or near 4 P. M. at the end of the school session, as they would have been had that group been given the work at 9.40 A. M. At least, when a class, all of whose members had been doing the work at say 11 A. M., was divided into two sections, one of which worked thereafter at 9.40 A. M. and the other at or near 4 P. M., the former at once did definitely better. In three or four days, however, their superiority had decreased markedly.

Heck ['13] has made an elaborate study of the efficiency

TABLE 18.

EFFICIENCY IN SCHOOL TASKS EARLY AND LATE IN THE SCHOOL SESSION. After King.

	9.30 A.M.	10.30 A.M.	11.30 P.M.	1.30 P.M.	2.30 P.M
ADDITION					
Columns done	2420	2341	2667	2359	2602
Mistakes	128	129	159	182	198
Relative efficiency, counting two columns off for each mistake, and expressing in terms of 100 as the average achievement	100	96	111	92	101½
MULTIPLICATION					
Amount done, each example counting 6	1313	1568	1664	1224	1577
Mistakes in calculation	144	195	238	188	285
Mistakes in method:					
Adding or subtracting	11	11	13	25	9
Partial product omitted	17	16	19	40	32
Failure to add partial products	2	..	1	..	2
Relative efficiency, counting 2 (i. e., one third of an example) off for each mistake, in calculation or in method	99½	116	115	75	95
JUDGMENT					
Relative efficiency	109	83	92	109	106
TOTAL RELATIVE EFFICIENCY	103	98	106	92	101

of school children in computation at four periods in the school session (9.10, 11.05, 1.10 and 2.30), using a combination of the methods of Bolton and the author. A sample of the four tests used is shown below. Each pupil was tested at each period, but with a different test. Of the 1153 pupils measured, one fourth took the tests in the order 9.10, 11.05, 1.20 and 2.30; one fourth, in the order 11.05, 1.10, 2.30 and 9.10; one fourth, in the order 1.10. 2.30, 9.10 and 11.05; one fourth in the order 2.30, 9.10, 11.05 and 1.10. Ten minutes was the time limit in each test. As the average result for the entire group he finds:

At 9.10 A.M., 140.37 units of work done, with 87.40% correct
" 11.05 A.M., 142.57 " " " " " 86.08% "
" 1.10 P.M., 142.67 " " " " " 86.17% "
" 2.30 P.M., 143.68 " " " " " 85.46% "

A second study by Heck ['13a] has confirmed these findings.

THIRD TEST

(1) ADD	(2) SUBTRACT	(3) MULTIPLY	(4) DIVIDE
123 9012 204 2030	946732 812021	1202 132	321)645852
(5) ADD	**(6) SUBTRACT**	**(7) MULTIPLY**	**(8) DIVIDE**
7403 465 5200 8001 2030 756	81014232 47637698	40658 564	436)3297904
(9) ADD	**(10) SUBTRACT**	**(11) MULTIPLY**	**(12) DIVIDE**
56996 76356 68665 59877 77563 58948 35869 75883	15264354 8678575	68597 876	976)7707472

Bergström ['94], Marsh ['06] and others have measured the efficiency of various mental functions at various times of the day. The general result of Bergström's work was that in sorting cards into piles according to various plans, the morning and afternoon work of six subjects did not differ in efficiency, but that in the evening about a tenth more time was required. In elaborate tests on himself from 8 A. M. to 6 P. M., Bergström found only trifling changes in reading numbers, adding and multiplying. In learning nonsense syllables, his afternoon work rate was about one twentieth below the morning rate, and another subject's was about one seventh longer.

The general result of Marsh's work is that in the functions tested (easy functions such as adding, marking A's on a page of printed capitals, sorting cards, or memorizing numbers) efficiency is about one fifteenth greater in the middle of the day and through the afternoon than in the early morning (7 to 9 A. M.), and about one thirtieth greater than in the late evening (9 to 11 P. M.). From the results so far obtained it is impossible to tell how uniformly the different functions follow this central drift. Different individuals seem not to follow it even approximately. It is also hard to tell how far the results measure an inner daily rhythm, inborn or acquired, and how far they measure such a rhythm complicated by the influence of a usual day's work with its warming up, fatigue, and suggestions from the habits of the individual and the community.*

With Marsh's subjects the tests were separated by the ordinary occupations of a student's day or by experimental work. Each of Bergström's subjects was asked "to be as regular as possible in his habits and to follow an average routine of work and rest." ['94, p. 249].

There is a remarkable unanimity in the results summarized in this section in showing that *ability* to work is, in school pupils, throughout and at the close of the school session, almost or quite unimpaired.† When the effects of both

* For *muscular* strength and skill, the afternoon does seem most favorable.

† Whether school pupils do, in fact, ordinarily achieve much less in late than in early periods is left as a question for educational experiments to decide. The common impression amongst teachers that they do may be to a large extent illusory.

novelty and practice are eliminated, no differences in achievement appear as a result of the work of the school session. Where they are not eliminated the former seems approximately to balance the latter. It is clear from the facts summarized that the assertions made in text-books on school hygiene that there are great and important differences between the results of tests at different periods of the school session, are quite unjustifiable. The very results referred to in support of these assertions disprove them.

If we take account of the conditions of the experiments and replace partial by adequate measures of efficiency, the apparent conflict between the results of Sikorski and Friedrich on the one hand, and those of Bolton and Thorndike on the other, turns out a harmony. The work in late periods is really only a little better or a little worse than in early periods. There is a general tendency of school children to increase speed at the expense of precision in repeated tests in adding, copying and the like. This fact has permitted the illusion of great deterioration to and the fallacy of interpreting a series like "99 right, 1 wrong; 98 right, 2 wrong; 97 right, 3 wrong" as "1 wrong; 2 wrong; 3 wrong," and then as "Early Efficiency = 1; Middle efficiency = ½; Late efficiency = ⅓." It is the false inferences only that are in conflict with the result, "Efficiency approximately equal at all three periods."

The two cases of deficiency of any large amount late in the school session are the results of Ebbinghaus with the completion of passages, and the results of Winch when pupils who had worked together at 11 A. M. were separated into divisions working early and late in the day.* The first of these cases is subject to all the sources of error due to the

* The work of Lobsien ['99] with the reading of unspaced mixed letters has not been reported here because the task is so evidently one of the control of the eyes, and because the results selected for his report may well represent the consequences of special habits in the two pupils concerned rather than any general facts of work and fatigue.

administration of the tests by class-teachers and is stated by Ebbinghaus himself to be unreliable. It is also in conflict with the results obtained in the same investigation by other methods. The second is ambiguous, since day by day the difference between the early and the late period grew less. It might well have been that in eight or ten days' tests the 'late' group would, by becoming adapted, have equaled the 'early' group. They seem to have been on the way toward such an event.

There are not enough studies of the effect of a full day's (eight to twelve hours) general intellectual work upon efficiency in special tests in the case of either adults or children. What there are show remarkably small differences between early and late records.

Winch's study with evening school pupils, which shows larger differences, deals with a different problem from either of those just described. Winch himself states clearly that it deals with a practical question in school administration, not the general question of the effect of mental work.

All these results are composites of ability to work and willingness to work, the latter often being greater than it would have been without the stimulation of the test, the stranger in the school-room, the novelty of the work and the like. What a person *can* do after certain work is not a measure of what he would ordinarily do thereafter. Here, as elsewhere, we have measured a composite of (1) the actual strength of certain bonds and (2) their readiness to act. The data show that the former is not seriously impaired, since there is no reason to believe that the course of school work makes pupils enough more satisfied by the work to compensate for a decreased strength of bonds. They do not show whether readiness is, or is not, diminished.

These facts do not, however, absolutely disprove the assertion that a diminution of efficiency in one function produced by its exercise without full rest will produce equal diminution in the efficiency of all others. For it is not, by the experiments themselves, proved that the functions exercised in school or in professional work did *themselves* lose in efficiency as the day's work progressed. That assertion, is seen, however, in the light of the facts of this and the last two chapters, to be surely false in one sense, and of relatively little practical import in the other. Diminution in what one actually does do certainly does not 'transfer' from a hated task pursued less and less zealously hour by hour to an enjoyable though difficult game taken up when the task is done or abandoned. No elaborate measurements are needed to support this obvious fact of every day. Diminution in what one can do in some one function, if upheld by sufficient motives and resolution, likewise probably does not transfer fully and equally to all other functions. Even if it did, the amount of exercise without rest that is required to produce even a moderate diminution in this sheer ability independent of zeal, is so great that, as a determining factor in actual work, transfer of the diminution is of little practical consequence.

THE SYMPTOMS OF MENTAL FATIGUE

Those who take it for granted that continuous mental work or even mental work with moderate pauses *must* diminish efficiency for any and all forms of work, naturally assume that the important duty of investigation is to find convenient symptoms of this general diminution of efficiency.

The only adequate way to test the value of such a proposed symptom—for instance, the ability to know whether one is touched by one point or by two points close together—would be to measure a subject's general diminution in efficiency on occasions when it was slight, when it was moderate, and when it was extreme; and to measure at the same time his condition in respect to the symptom. If the correlation between (a) the diminution in general efficiency and (b) the inability to discriminate one touch from two touches is perfect, the latter is an accurate symptom.

No proposed symptom has ever been so tested, and the literature of the subject, comprising reports of this, that or the other test made upon individuals in whom some general diminution in mental efficiency was suspected—but not proved, much less measured—is especially annoying. Very rarely have the writers even tested more than one symptom at a time, so as to measure their agreement. Almost never do they give data permitting an estimate of the reliability of their measures of the symptom itself. As a consequence, not much can be inferred from them, and I shall not trouble the reader with their details.

Tactual Discrimination.—The symptom which has been most studied is the increase of the distance at which two simultaneous touches on a given part of the skin must be in order to be known to be two. Griesbach ['95] asserted that school work, even of only an hour or two, caused marked loss in the power to distinguish two touches close together on the skin from a single touch; and the use of this symptom to infer fatigue is commonly called the Griesbach method.

Vannod ['96], Wagner ['98], Larguier des Bancels ['99], Baur ['02], Schuyten ['04], Sakaki ['05], Binet ['05], Noikow ['08], Abelson ['08] and others agree with Griesbach in finding vaguely that tactile sensibility is more obtuse after mental work than after rest. They disagree in respect to the amount of the change. Germann ['99], Leuba ['99], and Ellis and Shipe ['03] find that it is not. Bolton ['02] finds that it is not, in any uniform manner; and that increases in the amount of continuous work do not produce appropriate increases in the sensory obtuseness. Blazek's results ['99] are quoted by Griesbach as in agreement with his own, but, so far as I can see, seem to show no clear differences in sensory acuity before and after mental work. Griesbach ['05] has made a second study of the matter, criticizing Germann, Leuba and Bolton, and testing the method after a fashion. The test is by comparing the number of additions per minute at various periods of the day with the delicacy of discrimination of one from two touches.

Griesbach and his closer followers make the test of discrimination in approximately one minute. An accurate test requires thirty times as long. In some cases at least, they have thrown out the records of individuals who showed sensory obtuseness *before* mental work. Griesbach gets reductions of from ten to thirty per cent in the speed of addition from the work of the school day, which are almost certainly due to his unconscious suggestions to the subjects, since no other investigator has found anything of the sort.

The work of Germann, Leuba and Bolton is much more careful, and it is the writer's opinion that a subject who is induced to do his best at discriminating the number of touches and is left free from suggestion will show nothing like the obtuseness after mental work that Griesbach expects.

It is possible that both parties are right. Just because telling two touches from one is, for the ordinary man, a trivial matter in connection with which he lacks standards of achievement, he may permit natural slovenliness to operate in it after work, making it a test of inclination rather than ability. A poor record might then be symptomatic of fatigue in the sense of a readiness to indulge oneself by slack attentiveness. A scientific student or practiced subject, on the contrary, who regarded the task as one in which he could do very well or very badly according to the pains he took, and who did his best at each test, would show little or no greater obtuseness after mental work, since he would compensate for any impulse not to be careful. Griesbach's schoolboys, soldiers, engineers and the like might then give positive results as they did; and the practiced scientific students tested by Germann, Leuba and Bolton might for the same reason give negative results, as they did.

The Association Test.—Aschaffenburg ['99] found that in the course of a night of continuous work without food "the conceptual relations between the word given as stimulus and that given as response were gradually relaxed, and replaced by such forms of association as owe their arousal to long accustomed use; as a consequence, linguistic relations especially predominate. As exhaustion progresses, the associated ideas work less and less by their meaning; their sound and tone-color instead determine the response." [p. 81] His tests were made at approximately 9 P. M., Midnight, 3 A. M. and 6 A. M. The number of misperceived words rose; the time of response did not.

The Visibility-Invisibility Ratio.—If, on a white or gray background, a slightly grayer spot or stripe is painted, this becomes alternately visible and invisible to one steadily looking at it. Pillsbury ['03] studied the relations (1) of the length of the period when it was visible (V), (2) of the length of the period when it was invisible (I), and (3) of the magnitude of the ratio $\frac{V}{I}$, to the amount of work done (without much rest) in the case of six adult students. Three of these showed small differences, or none at all, between morning, noon, early afternoon and evening; one showed a very great increase in V and $\frac{V}{I}$ in the afternoon, and a still greater increase in the evening (with marked decrease in I). One did about the same at all periods save early afternoon, when he was inferior; one was markedly inferior in the evening, and in general after hard mental work.

Dynamometric Tests.—The force that can be exerted by the muscles has been proposed as a symptom of mental inefficiency due to work without rest. The facts seem to show that it is not (*e. g.*, Bolton, '02, Ellis and Shipe, '03, and Schuyten, '04), provided the subject works with full zeal.

The Rate of Tapping.—Wells ['08 a] has suggested the rate of tapping during 30 seconds as a symptom and measure of fatigue,* but not, as I understand him, of fatigue in the sense of a temporary diminution of general intellectual

* Especially in insane patients or others whose conscious coöperation cannot be counted upon.

efficiency due to work without rest, so much as of 'fatiguability' or the tendency not to bear up under the pressure of work. In any case the difference between the rate of tapping before, and that after, long hard continuous mental work has not been shown to parallel the difference in the efficiency of the work itself. He later finds ['08 b, p. 469 f.] that the fatigue of the one hand is not a close symptom of that of the other, either amongst individuals or amongst the different days of the same individual, and wisely remarks that this limits largely "the value of the single measure as a general measure."

The attempts to discover some convenient symptom and measure of fatigue seem misguided. If there is general fatigue, in the sense of a lowered level of mental energy, or any other one entity or condition caused by work done without rest and manifested equally in all functions, the condition of *any* function should be a symptom and measure of it. If there is no such thing, it is futile to seek its symptoms.

The reality that most nearly approximates such a general incapacity due to lack of rest is a condition known to common sense as inability to attend, 'concentrate,' or think against resistance. These words refer to a condition of intolerance of any effort and of inability to do the mental work demanded by schools, business, household or professional occupations without effort. The intolerance and inability are not absolute. Some favorite occupation, in itself as 'hard' as those avoided, may be sought; and, given a sufficient motive, effect will be made. But to the individual, any task such as is likely to be required of him, seems, and in a sense is, destined to secure an inefficient response. The symptoms of this condition could, by enough ingenuity and care, be discovered. The excessive aversion to work, especially to beginning work, is the only one patent to common observation.

The Feelings of Fatigue.—Teachers, physicians, employers, and mental workers themselves use oftenest the worker's so-called *'feelings of fatigue'* as a symptom of the degree of inefficiency.

Little is known about the relation between the feelings of mental fatigue and the fact of diminished mental efficiency. This is in part due to disagreement about what shall be included in the former. If individuals are allowed to decide which of the feelings that they have after mental work shall be called feelings of mental fatigue, they will report a great variety, including most frequently feelings of effort or strain and of aversion, but also irritation, worry, excitement, mental emptiness, confusion, weakness (expressed as 'Felt that I could do no more' or 'Felt no power to do anything'), sleepiness, pains in the eyes and head, and many others.

No careful attempt to discover the relation between these feelings, of whatever sort, and the inefficiency of any function or combination of functions has ever been made, but there can hardly be any close correspondence between the so-called feelings of fatigue and the degree of diminution in efficiency. In experiments with sixteen educated adults who worked at mental multiplication with three-place numbers, for from 3 to 12 hours, with no rest or only a brief period for luncheon, as a prescribed task in a graduate course in psychology, only slight correlation appeared between the intensity of the feelings of fatigue reported and the degree of fatigue itself. Amongst sixteen cases, one finds the following (each number under *Loss in Efficiency* represents the increase of time required at the end of a work-period over that required after abundant rest, expressed as a percentage):

Individual	Report of feelings of fatigue made by the individual	Loss in Efficiency
No. 3.	"The principal fatigue which I felt was physical; resulting from sitting so long in one position and the eye-strain."	197
No. 4.	"I think my mind was clearer at the close than at the beginning in both (series of tests)."	81
No. 9.	"I did not suffer any great strain in this test, though it was exceedingly tedious."	68
No. 10.	"I felt alternately tired and rested and rested as I worked; sometimes my head ached. At the end I did not feel specially tired."	36
No. 2.	"Very tired last hour."	36
No. 5.	"(Feeling of) fatigue was gradual until about the last hour. Then I completely gave out."	—9
No. 6.	"Being exceedingly dizzy and having headache at the end of two hours, I thought it best to discontinue."	—15
No. 7	"I was too tired for expression during the seventh (the last) hour."	14
No. 12.	"(At) 6 hrs. 32 mins. became utterly exhausted; could not hold numbers at all. Exhaustion came gradually."	83
No. 14	"From 3.05 to 4.27 I worked with the greatest effort. At that time I was more exhausted than I ever have been before. . . . Fearing that further application might be too much of an eye-strain, I stopped work at 4.27."	118

The individuals who felt most fatigued, judging from their reports, and those who felt least fatigued, did not differ appreciably in the actual loss of efficiency. Within the same individual's record there is a similar lack of anything like close correspondence between the rise of feelings of fatigue and the fall in efficiency.*

In ordinary life the relation is presumably closer, since the common response to such feelings at a certain intensity is to intermit work or stop it altogether. But the fact remains that if an individual does continue in spite of them, the intensity of his feelings of fatigue is a very, very inadequate measure of his loss in efficiency. It is an important fact that in an experiment in continued work in adding, memorizing, mentally multiplying, or the like, a subject *feels more*

* The data justifying this statement are not presented here.

and more fatigue, while his *work,* by the effect of practice, *becomes better and better.*

The feelings of fatigue, from what little is known of them, thus seem a very poor symptom of the loss of ability. They are not a mental ergometer, increasing in intensity in some regular relation to the amount of energy that is used, and waning steadily hour by hour during rest. They operate to attract one to work, to stop one from working, and to prevent one from renewing mental work, but they do not tell how badly one will work if he does continue as efficiently as he is able.*

THE RELATIONS OF 'MUSCULAR' WORK AND FATIGUE TO 'MENTAL' WORK AND FATIGUE

One of the misconceived problems arising from the supposition that all forms of mental work without rest produce alike some one condition—'fatigue of the mind' or 'general mental inefficiency—is that of the relations of mental fatigue and muscular fatigue. The problem has been put as that of a pair of simple problems: 'How far does mental fatigue produce muscular fatigue?' and 'How far does muscular fatigue produce mental fatigue?' But, obviously, there are very many relations, each to some extent independent of the others, and so, by rights, to be measured by itself.

The experimental work so far has neglected to make such analysis, and has in fact been confined almost exclusively to (1) the differences, early and late, in a school-day, in the force of the movement of squeezing a hand-dynamometer or of flexing a finger against resistance, and (2) the differences

* Lindley notes that the variations in speed of adding showed no uniform relations with subjects' reports as to "weather, health, power to work, apparent ease or difficulty of the daily task and the like." ['00, p. 485] Whipple ['10, pp. 262 and 265] found that, in approximately two hours' work in marking the letter *c* or *o*, the last half hour's work was 98 per cent as effective as the first half hour's work, in spite of the presence of "marked subjective fatigue."

in certain intellectual tests before and after some defined amount of physical exercise.

The first group of experiments will not be reported here, save for the statement that, when properly conducted, no differences of consequence are found.* There is, however, in many of these cases, no evidence of any loss in *mental* efficiency. I report briefly the results got by Bergström, Bettman, Miesemer and Hylan in connection with the second problem.

Bergstrom ['94] found that exciting physical exercise (tennis or ball-playing) from 4 to 5 P. M. increased this efficiency in reading numbers, adding, multiplying and memorizing nonsense syllables.

Bettman ['95] found, in one set of experiments, that a two-hours walk shortened his reaction time (with choice) by 13 per cent but at the cost of an enormous increase in the percentage of wrong reactions (2.6% to 26.9%). An hour's adding lengthened it by 31 per cent, but reduced the percentage of wrong reactions from 2.6 to 1.0. Another subject showed the same general tendency. Another series, with himself as subject, showed a slight lengthening by the walk with a not so great increase in wrong reactions, and a lengthening of 14 per cent from the adding with a reduction of wrong reactions from 1.6% to .5%. A third series showed much the same result as the second. The times after rest, walk, and adding respectively, were .309, .340 and .349 seconds; the percentages of wrong reactions were 3.6, 0 and 19.5. With word reactions his averages were .317, .357 and .391 seconds after rest, walk and adding respectively. Wrong reactions were not recorded, though Bettman gives some evidence to show that precision was least after the walk and greatest after the adding. In committing lists of digits to

* As has been noted, the most careful investigations, in which the factor of interest has been more or less controlled, show that the work of a school-day for a child [Schuyten, '04], hard study from 9 A. M. to 5 P. M. for older students [Ellis and Shipe, '03], and two hours addition at maximal ability by adults [Bolton, '02], all alike fail to lower muscular achievement of this sort.

memory Bettman did about seven tenths as much after the two-hours walk as after rest. The effect of the hour of adding was almost exactly the same. In adding Bettman did about 12 per cent less after either the walk or the adding than after rest.

Miesemer ['02] compared his efficiency in grasping and holding in mind groups of letters shown briefly, before and after an hour's rest, an hour's walk and an hour's adding. He found the effect of the walk inferior to that of the rest, and the effect of the adding to be the worst of the three. Seven per cent fewer letters were taken in, and one and a half per cent fewer of them were right, after the adding, than before; whereas, on the six rest-days, four per cent more letters were taken in (one per cent fewer of them being right) after the rest than before. Two per cent more letters were taken in (six per cent fewer being right) after the hour's walk than before. The same result held for tests in which the letters were held in mind for thirty seconds. After rest, there was a gain of fifteen and a half per cent in the letters reproduced, and of two and a half per cent in the percentage of them that were correct. After the adding, the number of letters reproduced dropped two and a half per cent and the percentage of correct ones four per cent; after the walk the number reproduced remained the same, but one and a half per cent fewer were correct.

From the data given by Hylan ['02], it appears that a two hours walk brought a slight improvement in his ability to add for twenty minutes (in four sections, separated by rests from 0 to 20 minutes long). On the whole, he performed 19,485 additions on the rest-days, and 20,146 on the days when the work followed a two hours' walk.

The general impression which these four studies together make is that surely there is no uniform effect of muscular work upon mental efficiency and that the average intrinsic effect of moderate amounts of it is very slight.

The main identical elements in so-called mental and so-called muscular fatigue in ordinary life seem to be the deprivations, the headaches and sleepiness. The impulses to familiar forms of enjoyment, whose inhibition the continuance of effective work requires, may be in large measure the same, though mental work is at a disadvantage in commonly depriving the worker of certain out-of-door indulgences which muscular work often permits. Headaches are more frequently a consequence of mental, and sleepiness of muscular, work, but either a headache or sleepiness may be the consequence of either sort of work and may increase the inner resistance to either sort of work.

CHAPTER V

General Theories of Mental Work and Fatigue

DEFINITIONS

We can now return to some of the fundamental facts about mental work and fatigue. In the course of our survey of the facts known concerning mental work and fatigue we have been led to define the real questions involved so as to relieve them from vagueness and to discourage merely verbal answers to them. Instead of pretending to describe changes in mental energy available following upon energy expended, we have measured the changes in the quantity and quality of certain products when the individual produces them as incessantly as he can. The change in, say, four hours of such production can be studied by itself, or be compared with the change in four hours of production distributed in any other way. The production at the end of such a period can be compared with that after an interval of no production.

We could also measure, though this has not been done, the satisfyingness or intolerability of the process of production at any stage.

Mental work has been noted as an ambiguous term, meaning on the one hand (1) mental achievement—the production of certain products, and on the other hand (2) mental effort—the initiation or continuance or prevention or cessation of a certain response in spite of the intrinsic relative unsatisfyingness of that behavior.

Mental rest is similarly ambiguous. In thinking about it one should make clear whether he is concerned with (1) mental inactivity—the absence of mental achievement, or with (2) mental relaxation—the absence of mental effort.

The *efficiency* of a function may be defined as the quantity and quality of the product produced (1) per unit of time, or (2) per unit of time with a given amount of effort. Call the former the *gross* efficiency of the function; and the latter, its *analyzed* efficiency.

If we stick to the first meaning—the objective fact—in each case, we have a series of useful objective definitions of important facts in behavior, which may be called work, rest and fatigue, as follows:

Mental work (achievement) is the behavior of an organism whereby certain products* are produced. Continuous mental work means the behavior of an individual who is producing as incessantly as he can.

Rest (inactivity) of a single function is an interval in which the individual does nothing toward that sort of production. General rest similarly would mean an interval in which he did nothing toward production of any sort. This perfect general rest is of course only approximated.

The fatigue of a function is that diminution in its productivity or gross efficiency which inactivity† can cure.

The efficiency of a function, except at its physiological limit, is always subject to a possible improving effect from exercise. Except when there has been overlearning beyond that required to bring it to the physiological limit, it is always subject to the law of disuse during rest. We have, however, seen that the effect of fifty or sixty hours of disuse upon such functions as adding, multiplying, reading, cancelling, typewriting, translating and the like is very slight.‡ Hence, neglecting the effect of disuse for less than fifty hours, the fatigue (F_x) of a function (in the above sense) that is due

* Such as poems memorized, books written, problems solved, decisions made, houses planned, lessons taught or prepared, and the like.

† Either of that function or in general as may be specified.

‡ Save in the case of the very early stages of this learning.

to any given amount (X) of continuous exercise (X being not over ten hours) may be conveniently measured as follows:

Let Y = an adequate length of rest.

Let P_{ry} = the productivity or gross efficiency of the function after y hours rest.

Let P_{ry+wx} = the productivity or gross efficiency of the function x hours later, the x hours being devoted to continuous exercise of the function.

Let $P_{ry+wx+ry}$ = the productivity or gross efficiency of the function after the x hours of exercise plus y hours of rest following.

Call the fatigue due to the x hours of exercise, F_x.

Then $F_x = P_{ry+wx+ry} - P_{ry+wx}$.

If the function is at the limit of practice, $F_x = P_{ry} - P_{ry+wx}$

The adequacy of the rest can be tested with functions at the limit of practice, any y being then adequate which brings P_{ry} and $P_{ry+wx+ry}$ to this limit. Or, in general, y is adequate when $P_{ry+wx+ry}$ is as great as $P_{rz+wx+rz}$, z being any longer time than y. In ordinary experiments, all that is necessary is to make y abundantly large, say, equal to 3 x or more.

F_x, for values of x from one to ten hours, has been shown to be very small. In general, under pressure from the determination to continue doing one's utmost, the associative mechanism of the brain involved in any given mental function seems to work for a long time with a very slight decrease in gross efficiency. But fatigue is not necessarily, and is probably not in fact, so slight, if it is measured by the diminution in *analyzed* efficiency—in the productivity per unit of time with a given amount of effort. It may well be that in order to maintain the same degree of *satisfyingness* at the end as at the beginning of the five hours of work, the individuals referred to in the above measures would have had to relax in speed and carefulness so much as to have shown a decrease in efficiency of 30 or 40 per cent on an absolute scale; or an increase in the time required, accuracy being constant, of over 100 per cent.

The great present need in experimentation on mental work is to measure mental effort as fully as mental achievement, and so to compute the changes in analyzed efficiency—the quality and quantity of the product per unit of time with a given amount of effort.

We may distinguish the maximum power of a man's neurones to make certain connections from their readiness to do so. A man may be *able* (in the former sense) to multiply 629 by 736 as quickly as ever, at a time when the work is ten times as intolerable.

If we were able to measure the satisfyingness or intolerability of the process of production at each stage we could give a more adequate account of mental dynamics. We might then define our facts more fundamentally as follows:

Let B_s be the group of bonds whose strength improves a given function, F.

Let B_w be the group of bonds whose weakness improves the given function.

Denote increasing degrees of *strength* of B_s by S_{so}, S_{s1}, S_{s2}, etc.; and of B_w, by S_{wo}, S_{w1}, S_{w2}, etc.

Denote increasing degrees of *readiness* of (or satisfyingness of the action of) B_s by R_{so}, R_{s1}, R_{s2}, etc.; and of B_w, by R_{wo}, R_{w1}, R_{w2}, etc. Let $R_{s(-1)}$, $R_{s(-2)}$, etc. and $R_{w(-3)}$, $R_{w(-4)}$, etc., denote degrees of *unreadiness*.

Then the temporary effect of any defined production, under any defined conditions, during any given time T, upon F, could be measured by the changes in the S's and R's of the B's in question. The effect upon the 'mind as a whole' would mean the effect upon the S's and R's of *all* B's.

Each statement of temporary deterioration by exercise or improvement by disuse would then specify the changes in both the mere strength of bonds and their readiness to act; and the majority of present mysteries, apparent contradictions and vaguenesses concerning mental dynamics would be on the road toward being cured. We would not need to talk vaguely in terms of fatigue, or temporary deterioration, but could speak of the strength or closeness of defined connections and of their readiness to act. We could, in all probability, then understand certain important facts—for example, that it is the growing readiness of competing bonds, (B_w's) rather than an intrinsic weakening of the B_s's that so often causes deficiency in work as it progresses without rest; and that, in the experimental results, the advantage of rests to permanent improvement is greater than their advantage to temporary efficiency.

All the facts, both of experimental studies and of everyday life, support the hypothesis that the effect of continuous exercise upon *readiness* is far quicker, greater and more significant than its effect upon *maximum power.* Fatigue in the vague popular sense means that we are less willing rather than that we are less able, that the probability of achievement is decreased by the increased effort that it requires rather than that the possibility of achievement is decreased inevitably,—that the activity of the function becomes less satisfying rather than intrinsically and necessarily feebler.

The ultimate physiological explanation of the phenomena of mental work and fatigue will therefore, I venture to prophesy, be found largely in the conditions of readiness and unreadiness of the neurones, and the main practical problem of the administration of mental work will be found to be the problem of interest.

DODGE'S THEORY OF MENTAL WORK

The only serious attempt to define and measure mental work in any matter of fact, objective way other than by the products produced is that of Dodge. In an important paper ['13] he attempts to show that 'amount of metabolism in the neurones' is the proper meaning of 'amount of mental work.' He says:

"Theoretically all nervous activity must involve metabolism. The fact of combustion, the use of oxygen, and the katabolism of nervous substance has been incontestably proven for the central nervous system of the frog. Even in the peripheral nerves, excitability and transmission are proven to depend on a normal supply of oxygen. Whatever attitude one may take towards the relation of nervous activities and consciousness, here at least is a promising direction of direct investigation. If fatigue is to mean anything to psychology more than an empirical fact of diminished production, it must correlate with the general physiology of complex nervous centers. But such correlations were absurd except on the assumption of a correlation between the general physiology of complex nervous centers and mental processes. And problems of general physiology are fundamentally problems of metabolism." ['13, p. 10]

His reasons for objecting to having work mean only production, and fatigue only "an empirical fact of diminished production" reduce essentially to the fact which I have been repeatedly emphasizing in these chapters, that the product produced per unit of time by a man under the same conditions of previous production and rest varies so enormously with the amount of accompanying interest, satisfyingness, effort, strain and the like; and, on the other hand, may be identical, though the amount of strain, effort, intolerability and later unfitness to produce which it brings vary enormously. He says:

"The common measure of mental work is number and time. In the adding test, for instance, mental work is measured by the number per minute. In reaction tests it is measured by the latent time. Fundamentally both measures are the same. It is a matter of expediency whether one measures a series of events so many per minute, or whether one measures the duration of each individual event. This output conception of mental work is convenient; practically significant, since it may deal with the actual processes of business and school; and it presents plausible physical analogies. It is the yardstick of mental achievement. If one can make *n* calculations per hour, *n* calculations is a plausible measure of the mental work of an hour. Decrease in the value of *n* is commonly supposed to indicate a loss of efficiency in the mechanism. By interpolation different tasks could apparently be correlated. Five sums in arithmetic are equal to ten pages of Dickens or half a page of German prose. We do practically measure off work that way. Hours of lecture are units for instructors; hours of recitation for students. If we would be extra particular we go back to the hours of preparation. For any individual student we balance off subjects in a similar way. The Latin lesson is twice as hard as the French lesson, etc. The conception is useful. Unfortunately, however, it takes absolutely no account of the conditions of work or the energy involved. It is essentially non-dynamic. To be usable it must be constantly assumed that one is doing his best, that the output is maximum or that the effort is sustained and equal. Unfortunately one reaches his maximum output perhaps only once in a lifetime. No one can do his best possible for an hour at a time. It is notorious that the ergographic curve depends not on capacity but on incentive, interest, and impulse. Our entire experimental knowledge of mental fatigue is on the yardstick basis. Time is a good enough measure for commercial and practical purposes, but an hour of lecturing to a class of one hundred and an hour of lecturing to a class of ten are somehow different in practical experience. An hour of experimentation and an hour of reading are different from an hour of bracing up a delinquent student. All are work but there is no use trying to give them scientific value on the basis of time. In factory terms one must ask how much of the plant is going under these different circumstances. The output is no real measure of work, though under some circumstances it might be a good measure of the efficiency of work. In no case does the directly observable output give any adequate measure of the dynamic process.

"It seems to me to be a serious common fault in all the familiar fatigue experiments that we have developed no adequate technique for controlling either the inner stimulus, the degree of effort and incentive, or the concurrent dynamic processes. We commonly measure only one more or less arbitrarily selected partial process on which instructions center. But instructions can guarantee sustained and equal effort only so far as there are reliable subjective evidences that effort has lapsed. If the subject is allowed to see the objective results, they introduce a totally new feature into the mental complex. They may constitute stimuli either for increased effort, or for discouragement and indifference.

"Products would be an adequate measure of work only when we could assume that the factory processes remain constant. But it is the experimental intent in fatigue that the processes shall change. What is it that actually does change, in any given case, the changing products do not indicate.

"Concurrent feelings of strain, effort, and exhaustion represent the noticeable organic effects of work. In many cases these introspective organic indicators are very delicate and useful. This is peculiarly true in active touch. But they seem to have no immediate relation either to the efficiency or to the amount of mental activity. It is notorious that introspective estimation of reaction and memory, of the speed and accuracy of the familiar mental tests is unreliable under the most favorable circumstances. With supposedly decreased mental capacity of fatigue, judgment should become even less accurate, and strains will change their meaning." ['13, pp. 7-9]

This admirable statement of the inadequacy of a measure of 'product produced in time T' to measure the effort required to produce it or the effects upon the organism due to its being produced seems to me, however, to be an argument for measuring these latter separately rather than for trying to sum them all up in a measure of the amount of metabolism which occurred in the neurones during time T, or some defined length of time thereafter. In order for measurements of the amount of metabolism in the neurones to be of any theoretical or practical use, they must be related to amounts of product produced, annoyance felt, repugnance overcome, effort made, bonds strengthened or weakened, sensitivities heightened or lowered, readinesses to conduct increased or decreased, or some known facts of behavior. We should need studies of the decrease in the externally observed efficiency of a mental function and of the changes in the satisfyingness of the process of production *all the more* if we had, along with them, measures of the metabolism in the brain as a whole and in separate neurone-chains within it. For the relation between these changes and the metabolism that occurred would be even more important than the relations between them and the ill-defined external and internal states of affairs which they accompany. Yet these latter relations, Dodge admits, are instructive. So also, if we could relate rests of various sorts to the metabolic processes which occur during them, we should have all the more reason for relating these various rests also to changes in power to produce a given product, and in the satisfyingness of producing it.

We may, then, welcome any attempts to measure the metabolism of any group of neurones under any defined conditions, not because this measurement will render measurements of the empirical facts of diminished production or diminished satisfyingness unnecessary, but because it will magnify their value. So far as verbal usage goes, whether we use the term *Mental Work* to refer to the production of defined products, to the overcoming of intolerability, to a combination of these, or to metabolism within the connection-system, is a matter of little consequence so long as nothing is inferred from the name.

THE MECHANICAL OR ENERGY THEORY AND THE BIOLOGICAL OR RESPONSE THEORY

In the early discussions of mental work and fatigue, the use of the term *Work* led thinkers naturally enough to follow the train of thought suggested by physics and to conceive of mental work as the consequence of expenditure of mental energy, of fatigue as the consumption of a stock of potential energy, and of rest as an opportunity for its restoration. If left vague enough, such a mechanical theory of the operation of mental functions does no great harm, but it is almost always misleading; and, when it is at all rigorously defined, it becomes, I think, either meaningless or wrong. Some of the reasons for preferring what may be called the Biological Theories or Response Theories or Extrinsic Theories of mental work and fatigue, may be briefly mentioned.

The first reason is that the rate of change in efficiency as more and more work is done without rest is not such as should be the case by a mechanical theory. It is far too irregular. The curves of work, special or general, have no such evenness as the curve for the pressure from a reservoir whence water runs out faster than it runs in, or the curve of force of impact of a ball dropped from steadily decreasing heights.

The mechanical theories of work consequently have to invent various subsidiary forces to act in conjunction with the loss of energy so as to produce the irregular course which efficiency actually takes under continuous work. For example, the fact of a gain in efficiency in the first ten or twenty or even forty minutes of work, contrary, of course, to expectation from the loss-of-energy doctrine, is attributed to the influence of Incitement or Warming-Up or Anregung. Similarly the fact of a frequent gain in efficiency in the last ten minutes of work, provided the worker is aware that they are the last, is attributed to a tendency to Final Spurt—to an increased 'exertion of the will' due to knowledge that the end of work is near. The resort to these subsidiary factors is, of course, an admission that loss in energy is not an adequate cause of the changes in the amount of work done hour by hour as work is continued.

The second argument concerns the enormous potency of interest in maintaining, and of repugnance in diminishing, efficiency in work without rest. Consider, for example, the effect of an offer, made at the end of the tenth hour of work, to give the worker a thousand dollars for every one per cent of improvement above his last hour's score. Consider similarly the probable work-curves resulting when a devotee of the game plays chess, and when he answers undesired and unprofitable questions, in each case for five or six continuous hours. Interest does not add to, nor does repugnance subtract from, a store of energy. By the mechanical theories rest and work monopolize these two functions. Interest could, at the most, only release the energy faster. But it is a fact easily verifiable that interest *does* add to, and that repugnance *does* subtract from, the amount of work done. The amount of work done then cannot depend closely upon the magnitude of a supply of mechanically conceived energy.

Finally the nature of mental work and of decreased efficiency in it, make the hypothesis of a usable and restorable supply of energy inappropriate. Consider any representative samples of mental work—*e. g.*, addition, solving geometrical problems, writing essays, devising arguments, correcting examination papers, reading proof. The work is the production of the *right* responses to certain situations. The mere amount of movement, of consciousness, or of neurone action is irrelevant. If we are to have a physical metaphor to illustrate mental work, we may say that the work of adding 7 and 9 is not like moving a pound through a foot against gravity, but is like the work of moving a pound of lead from a given space in Boston to a given space in New York. The mere physical work of the latter varies enormously, according to the condition of the vehicle used, the condition of the roads travelled, the route taken, and the opposition encountered from fire, flood, living animals and other natural forces. There is always a *qualitative* demand and a variety of obstacles to be overcome, and a choice of ways and means.

No physical metaphor is desirable. All that is meant by mental work is getting the required responses to certain situations. All that is meant by fatigue is the temporary diminution in the efficiency in making such required connections which comes from incessantly making them. Why the diminished efficiency should be so caused is a matter for investigation, not presupposition. All that a supply of 'mental energy' could properly mean would be a supply of power to make the required connections; and since what hinders making a connection in learning is its consequences, the reasonable expectation is that what will hinder making it in fatigue will be its consequences. An animal tends to repeat a connection when repeating it brings a satisfying state of affairs, and may be expected to discontinue it when repeating it annoys him. An animal would seem likely to discontinue or decrease mental work because continuing it annoys him rather than because some inner fund of impulsion, which might be likened to physical potential energy, was running low. The more promising theory would seem to be one that explained why mental work continued without rest became less and less satisfying.

This the Biological or Response Theory tries to do. Work without rest, it maintains, becomes less satisfying (1) by losing the zest of novelty, (2) by producing ennui, a certain intellectual nausea, sensory pains and even headache, and (3) by imposing certain deprivations—for instance, from physical exercise, social intercourse, or sleep.

That these facts of behavior are found where diminished efficiency as a result of work without rest is found, is a fact subject to verification by observation and experiment. Even the advocates of a mechanical theory will hardly deny it. That they cause the loss in efficiency is shown by the gain which follows their elimination. Varying the superficial form of arithmetical drills, while exercising the same mental function, will postpone the loss in efficiency by maintaining the force of novelty. The addition of a money reward, or of a demonstration that the work is useful for some desired end, or of competition for excellence, may temporarily abolish fatigue by abolishing the ennui. The common phrase that one is 'tired of' certain work represents a certain stage of fatigue better than 'tired by' it does.

The extreme condition where the mind seems literally nauseated—will not have anything to do with the problem—may be cured similarly by an increase in the value of the answers to be got. As a fatigued muscle can be given renewed efficiency by washing out or counteracting the products created by its action, so a fatigued mind can be in part restored by washing out ennui and repugnance by inattention or counteracting them by interest and motive. It is harder to eliminate experimentally sensory pains and headaches, but it seems probable that if these incubi could be lifted off, efficiency would rise.

That eliminating the deprivations, or in clearer phrase, permitting the indulgence of certain impulses, increases the efficiency of work is almost a crucial experiment for decision between the two classes of theory. When a boy regains efficiency by being allowed to walk up and down the room, or when the presence of a friend to study with her doubles a girl's achievement, it is clear that the previous deficiency was but little due to a lower pressure from a lessened reservoir of energy.

The effect of mental work without rest in causing deprivations, and of rest in permitting the corresponding indulgences, has been little studied. Attention has been centred upon what happens in the function that is working in disregard of the other functions which are being denied exercise. It is the fact that we are fatigued by what we do not do as truly, and perhaps as much, as by what we do. For children *not* to run and jump and squirm and sing and laugh and talk is the essence of mental work. For us all *not* to indulge in our favorite occupations is, as hour after hour of reading legal reports, or adding columns, or whatever the task may be, progresses, a more and more impressive feature of the task. Cases of special theoretic interest are those where the deprivation is from opportunity to do other mental work. For, in some such cases, the other work, deprivation from which fatigues, and exercises at which rests, the individual, would be rated by men in general as very exhausting. By the ordinary energy-theories it would involve large expenditures of mental energy.

If one could count up all the cases where individuals have stopped mental work and could know the chief cause in each case, it seems likely that the plea of some contrary impulse for gratification, some game to be played, sensory pleasure to be enjoyed, or the like, would be by far the commonest cause. Rest, again, except when spent in sleep, is not as a rule devoted to replenishing lost mental energy. It is far oftener devoted to indulging wants which mental work proscribes. To read, to talk with one's family and friends, to hunt or fish, to play active or sedentary games, and to make or listen to music, are occupations that often require a large expense of 'mental energy,' however defined, and that almost never approximate to the mental inaction of *dolce far niente* or sleep. They rest us by relief from strain and irritation, but not by cessation of mental action.

No theory of mental work and fatigue should then fail to take account of what continued work prevents the worker from doing. The little child who complained "I am tired of not playing," expressed admirably one feature of fatigue. The strain of not giving way to certain tendencies to response is as important as the strain of continuing certain others. Work in the popular sense is distinguished from play or recreation less by the amount of positive action than by the amount of restriction. We are fatigued by what we do *not* do.

On the whole, the biological theory seems much more probable. The effect of continuous mental work may be in part to use up some store of a complex of patience, self-control, vigor and the like, which may be called mental energy, but it surely is to produce certain annoying states to which the natural response is a diminution or cessation of the activity which causes them.

The behavior which results in certain products such as sums done, dictations written, paragraphs translated, and the like, is subject to the laws of all behavior, and to no others. If a continuance of the productive responses at the same speed and in such a form as to give equal quality, is satisfying to the individual concerned, he will continue them. If such continuance brings discomfort, he will tend to stop them altogether, or to intermit them, or to make them in such altered form and speed as lets them bring relative satisfaction. Stopping the work outright does not of course occur in the great majority of experimental investigations of fatigue, but is very common in ordinary mental work. Intermitting the work, dropping it, taking it up as thoughts of rewards, punishments, duty and the like, make idleness even more discomforting than the work, dropping it again, and so on, are also, in the nature of the case, rare in the experimental studies, but very common everywhere else. Relaxing speed and care and tension to such a degree that the work is less annoying than is the condition of not working (with the consequences attached thereto) is the device to which the subjects of the experiments are restricted. Whether one relaxes, intermits or stops work, the immediate reason is not that he has not the 'energy' to go on with it, but that he feels more comfortable to relax, intermit or stop it. Whatever parallel to a decreased store of energy there is, is effective chiefly by making the responses concerned in production less satisfying than they were before.

CHAPTER VI

The Hygiene of Mental Work

Readers of this chapter should remember that we are dealing with mental work—the work of the connection-system—not with either the work of the sense organs or the work of the muscles which so often accompanies it. This matter is of special importance in the case of fatigue of the sensory and motor apparatus of the eyes. So-called "mental" work in schools, business and professional life involves reading, writing or visual examination of objects to such an extent that the diminution of efficiency below what is desirable and the injuries from work are to a very large extent due to inability of the eyes to do what the mind requires and to overstrain of the eyes in the mind's service. It is well to keep sharply apart the means of increasing the efficiency of, and preventing injury from, purely mental work and the means of increasing the efficiency of, and preventing injury from, the use of the eyes. Interest, for example, rarely injures the mind, but may lead to very great harm to the eyes. Rest, in the sense of inactivity—the absence of any set task—may, through worry, depress or irritate the mind, but it is almost always good for the eyes. The theoretical and practical problems connected with the use of the eyes should form an important topic in educational hygiene, but in this book I shall not discuss them.

The practical application of the facts about mental fatigue may best be considered under the two topics—Desirable Means of Increasing Efficiency, and Desirable Means of Preventing Injury from Over-work.

MEANS OF INCREASING MENTAL EFFICIENCY

Roughly we may: (1) increase the organism's mental vigor or tendency to mental activity; (2) decrease the resistance, the forces inhibiting work; (3) improve the direction and method of activity; and (4) relieve the mind from the waste of excitement and worry.

The inner responsiveness of an animal to occasions for mental work is most economically improved by improving its general health. Other more direct influences limited to the connection-system there may be, but the safest hope is the maintenance of the health of the entire bodily machine. Consider the abolition of the effects of indigestion, rickets, chorea and scarlet fever, or of insufficient oxygen, food and sleep, in the case of children; consider the abolition of the effects of malaria, tuberculosis and alcoholism in the case of adults; consider even such a very minor factor as the common 'cold.'

The resistance which blocks mental work may be diminished by supplying interest and motive. It has been shown that certain kinds and amounts of mental activity are maintained without external subsidies, but much of what has to be done creates in the doing ennui, repulsion and pain, and deprives the worker of various satisfiers. The worker is thereby impelled to decrease, intermit or abandon the work. The resistances thus caused are not, however, inevitable, and curable only by rest. The same work done with interest does not so soon produce ennui and repugnance. The denial of certain satisfiers, such as games, conversation or reverie, may be balanced by the addition of new ones, such as a money-reward, zeal to improve, or confidence that the work will profit oneself and others. The inventor, man of science or poet, working a score of hours without rest at full efficiency, is not an exception to the laws of work, but an illustration of them. The limit of work for every man is elastic at the pull of interest and personal profit.

As a muscle becomes anew responsive to the stimulus, when the toxic products of its contraction are washed out or neutralized, so a mental function may be made to continue its output by washing out the repugnance and need for effort by an interest, and neutralizing the pain of restraint by a motive. In the case of wise and experienced adults, it is often hard to thus dissolve fatigue by adding interests and motives; they perhaps have already themselves used all the available ones. But the rank and file have not thus exhausted the preventives of repugnance and distraction; and children have hardly learned to use any of them. The children of a school class may work with doubled efficiency simply from learning the significance of the work to their wants, and associating the work with sociability, cheerfulness and achievement.

Since individuals differ in their interests, the proper distribution of the different pieces of work to be performed in the world will by diminishing resistance make the sum total done larger. If each man did the mental work for which he was fit and which he enjoyed, men would work willingly much longer than they now do. But if each worked only at tasks of real value and with the guidance of exact science, men could probably attain equal results though working far less than they now do. The best means of increasing efficiency are very simple ones—ceasing to learn by roundabout and stupid methods what is not so and ceasing to prepare with anxiety and pain for what will not occur. The time and effort wasted upon superstitions, pedantries and fads of which the science of the future will convict us, doubtless make the major part of our present burden.

One can hardly overestimate the value of peace and equanimity as means of increasing mental efficiency. Since nothing is done by worry or excitement that cannot be done better in their absence, there is nothing but gain in saving for achievement the time and strength now spent in ferment and ebullition. Too much of the life of home, school, industry, business, and even the professions, is still on a par with the war dances of primitive man. We need not burn down a house to roast a pig.

One of the problems of the economical management of mental work, that of the best arrangement of rest periods, has been somewhat elaborately treated by Kraepelin and his students. Their experiments were designed to measure the effect of pauses of different lengths, inserted after work periods of different lengths, upon the amount of fatigue. But they can be used to measure also the effect of subtracting for rest various portions of a given amount of time. Their results do not, however, give secure answers to either question. The work done (mainly the addition of one-place numbers, beginning anew when 100 is reached) is too simple and too easy; the efficiency is measured by speed only, regardless of errors; no measures of the degree of interest are given.

A report of some of the results will, however, make the problems themselves clear and will illustrate the difficulties of determining the effect of a rest.

Amberg ['95] measured his ability in adding (after the method described) on certain days, for an hour without rest, and on other days for two half-hours broken by a 5- or 15-minute rest, and also for twelve 5-minute periods alternating with 5-minute rests. The '30-minute work, 15 rest, 30 work' experiment was made by another individual also. Two subjects tested the '30-minute work, 15 rest, 30 work' arrangement when the work was learning series of digits.

Amberg also added for two hours separated by a 15-minute rest on some days, and with no rest on others. Finally, he had twelve subjects add for 60 minutes without rest on one day, and on another with three 5-minute pauses interspersed. The results were as follows:

The plan being '30 minutes adding, 5 minutes rest, 30 minutes adding,' Subject A's achievement after rest was on two days more (+ 4.7% and + 9.6%) and on two days less (— 9.6% and — 7.8%) than in the half hour before rest. On the days with no rest, the second half hour's amount varied from that of the first by + 0.4%, — 1.2% —4.8% and —5.3%. The median deficiency of work after rest was thus 1.55%; and after no rest, 3%. The median advantage due to the rest was thus 1½%. There are about 4 chances out of 10 that, in an adequate number of experiments, the median advantage of the rest to subject A would be 0 or less.

The plan being '30 minutes adding; 15 minutes rest, 30 minutes adding,' subject A's amount done after rest varied from that done in the first half-hour by + 1%, + 2%, +4%, and + 2% (+ 1.1, + 1.8, + 3.7, and + 1.9). On interspersed days with no rest, the amount done in the second half-hour varied from that in the first by + 24%, — 2%, + 1% and + 8% (+ 24.0, — 1.8, + 0.9, and + 8.0). For Subject B, in a similar experiment of eight days, the results were: Change after rest, + 10%, + 2%, 2%, and +4% (+ 9.8, + 2.4, + 1.9 and + 4.4); Change on interspersed days with no rest, + 28%, + 13%, + 1% and +7% (+.28.0, + 12.8, + 0.9 and + 6.9). The median increase after rest was thus, for A, 2% (1.85); and for B, 3% (3.4). The median increases after no rest were 4½% and 10% (4.45 and 9.8). The median injury of the rest was 2½% for A, and 6.5% for B.

The plan being '30 minutes memorizing figures, 15 minutes rest, 30 minutes memorizing figures,' Subject A's achievement after rest varied from that before rest by + 14%, — 2%, + 6% and + 2% (+ 13.7, — 1.8, + 5.7 and + 2.0). On four days with no rest, the number of digits learned in the second half-hour varied from that for the first half-hour by +6%, —2%, —26% and —18% (+6.2, —2.1, —25.9 and — 18.3). For Subject C corresponding figures are: With the pause, the second half-hour resulted in changes from the first half-hour of — 6%, — 40%, + 9%, + 15%, — 16%, and — 8% (— 5.8 — 40.4, + 9.2, + 15.2, — 16.4 and —8.1). Without the pause, the changes of the second from the first half-hour's achievement were —35%, +1%, —25%, —20%, and +9% (—34.5, +1.2, —25.2, —19.8 and

+9.2%). After the rest A showed a median gain of 4% (+3.9); and C, a median loss of 7% (—6.95). After no rest, A showed a median loss of 10% (—10.2); and C, a median loss of 20% (—19.8). The median advantage of the rest was thus 14.1 for A and 12.8 for C.

Taking the experiments in addition and those in memorizing both into account, the median effect of the 15-minute pause for all subjects is an advantage of three per cent.

The twelve subjects gained on the average 7%, 7% and 3% (+6.9, +6.9 and +2.6) over the first quarter-hour's amount in the three quarter-hours following, when three rests of five minutes each intervened. With no rests, the corresponding gains were 9.6%, 11.5% and 13.5%. There is, however, a discount to be made in the latter case, due to probability of a greater practice effect in the early course of an experiment, so that the average apparent injury due to the pauses is probably not over 3% or 4%.

The plan being 'five minutes adding, five minutes rest, twelve times repeated,' the amount done on four days was 6 %(+5.8) more than on four alternating days with no rest. The average advantage of the rest was probably not over 3%. Amberg estimates it to be 2.2%.

The plan being '60 minutes adding, 15 minutes rest, 60 minutes adding,' Subject A showed after rest differences of +5%, +3%, +1½% and —1% (+4.9, +3.0, +1.5 and —1.3). After no rest, the changes were +6%, —1%, —2% and 0% (+5.9, —1.4, —1.7 and —0.3). The median gain after rest was 2% (2.25); and the median loss after no rest, 1% (—0.85). The median advantage of the rest was thus 3%.

The results so far show that, for speed in addition of the sort described and in memorizing digits, the 5-minute pause, the 15-minute pause, and the 5-minute pause out of every 10, were all a trifle better than none. In the case of twelve subjects who added, on one day for an hour without rest, and on another for four 15-minute periods with 3-minute pauses interspersed, the rests had a slightly adverse effect. A 15-minute pause between two full hours of addition had, for Amberg, a trifling advantage.

The important lesson to be drawn from Amberg's results is that from small differences, each unreliable, nothing should be concluded without taking into consideration also the general probabilities due to general knowledge of human behavior in mental work. Amberg unhesitatingly accepts every difference, no matter how unreliable and unlikely, on general grounds, and concludes that, "In an hour's adding, pauses of 5-minutes in general exercise a slight but favorable influence upon achievement, while pauses of 15 minutes under similar conditions work decidedly unfavorably. So soon, however, as the work is extended to two hours, the effect of a 15-minute pause is reversed; it becomes favorable, whereas it was unfavorable in the case of a single hour's work. On the other hand we see the 15-minute pause act favorably in the case of memorizing even if the work-period is only an hour." ['96, p. 371] As might have been prophesied, later studies do not justify these sweeping conclusions.

Lindley, for instance, ['00] had three persons work at adding for two half-hours on each of 26 days, 6 days with no rest, and with rests of 5, 15, 30 and 60 minutes on 5 days* each. The changes from the first to the second half-hour were in percentages as shown below:

Individual	0 rest	5 min. rest	15 min. rest	30 min. rest	60 min. rest
A	+1.8	+0.8	+4.8	+4.6	+3.5
B	+4.2	+1.2	+2.7	+3.1	+2.5
C	—11.3	—7.0	—6.4	—3.9	—0.5
Average	—1.8	—1.7	+ .4	+1.3	+1.8

These experiments and others like them which could be quoted seem a rather paltry beginning in the study of the best possible distribution of time for temporary efficiency and permanent improvement, in schools, shops and homes. But they are better than nothing. The fact that the best number

* Four days for one subject.

and length of rests to be interspersed in sixty minutes of work are not yet known even for so simple a case as the addition of one-place numbers shows how much we have to do in the experimental study of the conditions of study and productive labor.*

A start has also been made at investigations of the very important question of the merits of giving up a long period to one kind of work compared with dividing it among several, by Weygandt. His study, 'On the Influence of Change of Work upon Continuous Mental Work' [97], should be repeated by other experimenters with many more subjects. It is very doubtful whether his measurements represent the eventual findings in the case of even the limited problem which he attacked. I shall nevertheless present them in some detail because, though questionable, they are the best information about the topic as yet offered by experimental psychology.

I have summarized the essential facts of Weygandt's forty-odd tables in Table 21. Each entry of it represents a rather complicated relation, which will be made clear by the following explanation:

Weygandt's general plan was to work 75 minutes continuously at some task (say memorizing lists of numbers) on two days (say, the 1st and 3rd of November); and on alternate days (say, the 2nd and 4th of November) to work 75 minutes, the first 30 and the last 15 at the task in question and the intervening 30 at some different task (say, adding). His question was: "How much, if anything, does the efficiency in memorizing numbers gain from the second procedure?"

If the functions in question had been free from any practice effect, he could have answered this by comparing the last 15 minutes work on the 1st and 3rd days with that on

* A body of facts, of great interest but very hard to interpret, exists concerning the experiences of employers of labor in shortening the working day and in enforcing a regimen of tasks and rests under the guidance of "scientific management." The most convenient single account of these facts is given in Goldmark's "Fatigue and Efficiency." ['12]

the 2nd and 4th. But he found commonly a great practice effect. For instance, in the given case the results were:

Figures Memorized in each 15-Minutes.

Nov. 1, (continuous work)	197,	269,	250,	310,	238
Nov. 2, (change)	360,	395,	....	....*	408
Nov. 3, (continuous work)	680,	636,	663,	541,	415
Nov. 4, (change)	864,	780,	....	....*	672

This influence of gain from practice, and also that of chance differences in the person's efficiency on different days, may be avoided by comparing the work of the last fifteen minutes with the work of the first thirty minutes in the two cases. Thus, from the record above, we find that with continuous work the subject learned 102 per cent (on Nov. 1) and 63 per cent (on Nov. 3) as many figures per minute in the last 15-minutes as in the first 30-minutes. Combining the percentages for each pair of days, we find the days of unchanged work showing 82½ per cent, and those of changed work showing 90 per cent. That is, the falling off of the last 15-minutes over the first 30-minutes (per minute) was 7½ per cent less when the replacement by addition was made than when it was not.

Figures corresponding to this 82½ and this 90 are given in the first two columns of Table 21. Thus the given case appears about half way down the table as: "A, Memorizing digits. . . . Addition . . . 82.5, 90," etc., and should be read, "On subject A, the effect of the continuous work with memorizing digits was a falling off to 82.5 of the first half-hour's rate; the effect of replacing the second 30-minutes of memorizing by adding was a falling off to 90 per cent of the first half-hour's rate.

Weygandt himself uses a still more complicated way of estimating substantially the same relations. He estimates what the subject would have done in the four 15-minute periods in question, if there had been no 'fatigue' whatever—if, that is, the effect of practice had been allowed to appear

* The 30 minutes being filled by adding.

in full unreduced by the effect of work done without rest. He then finds what per cent of this was accomplished in the two cases of uniform and of changed work. In the given case he estimates that, apart from 'fatigue,' the 238, 408, 415 and 672 would have been 387½, 532, 812½ and 975½, of which the actual numbers of figures learned were 61.4, 76.7, 51.1 and 68.8 per cents, respectively. With continuous memorizing, then, there is, as an average of the two days, a falling off to 56 per cent of the expectation with an unreduced practice effect. With the change to addition, the falling off is only to 78 per cent thereof.

Figures corresponding to his 56 and 78 are given in the third and fourth columns of Table 21.

Weygandt thinks that these results show that a change of work is less fatiguing than continuance of the same work only in so far as the change is to easier work. This is certainly not demonstrated by his records, and is probably not true. Of course, it is probable that, *other things being equal,* the easier the replacing work, the less will be the fatigue caused by it. But also, perhaps, other things being equal, the more *unlike* the replacing work is to the replaced, the less will be the fatigue caused by it.

To draw any conclusions concerning the differences between the effects of different rotations of work, one should repeat many such experiments as Weygandt's upon many individuals, each on many occasions. The only one of the combinations which he tried whose effect he can fairly be said to have measured, is the replacing of *Reading Hungarian* by *Adding*. Since the two experiments for each subject do not differ very greatly, and since four subjects were examined, one has a right to say that there is no considerable chance that the fifth quarter-hour of reading suffers very greatly either from the continuance of the reading or from replacing the second half-hour of it by addition; and that there is no considerable chance that one of these courses will be much more advantageous than the other.

TABLE 21.

THE EFFECT OF CHANGE OF WORK UPON EFFICIENCY. Computed from data of Weygandt, '97.

Subject	Order of Experiment	Fundamental Work whose efficiency was measured, without and with replacement.	Work by which the fundmental work was replaced.	Percent which rate of work in last 15 min. was of rate of work in first 30 min.		Percent which work in last 15 min. was of what it would have been by unreduced practice effect. Estimated by Weygandt.	
				Work being unchanged	Work being changed	Work being unchanged	Work being changed
A	2	Addition	Memorizing digits	86.5	72.5	83	69
A	8	"	Memorizing nonsense syllables	81.5	78	82	78
A	5	"	Marking a given letter	66	98.5	73	106
B	17	"	Reading an unknown language	81	107.5	68.5	92
D	19	"	Reading an unknown language	99	100.5	85	92
A	10	Reading an unknown language	Memorizing nonsense syllables	92.5	85	90	83
A	12	"	Writing at speed	86	91.5	84	89
A	11	"	Addition	95	95.5	93	94
B	16	"	Addition	104.5	102.5	97	94
C	18	"	Addition	94.5	97	83	88
E	20	"	Addition	94.5	80	93	79
A	14	"	Reading a second unknown language	88	101	79	90
A	15	"	Reading a third unknown language	95.5	86	87	78
A	1	Memorizing digits	Addition	82.5	90	56	78
A	6	Memorizing digits	Marking a given letter	58.5	59.5	53	54.5
A	9	Memorizing nonsense syllables	Addition	69.5	137.5	67	132
A	7	Memorizing nonsense syllables	Writing slowly	63.5	118	60	108.5
F	22	Memorizing nonsense syllables	Reading a fourth unknown language	110.5	74.5	110	74.5
A	3	Marking a given letter	Addition	69.5	65.5	65.5	62
A	4	Marking a given letter	Memorizing digits	52.5	108.5	53	108
F	21	Reading a fourth unknown language	Memorizing nonsense syllables	102.5	111	93	102

It is to be hoped that all of Weygandt's experiments will be soon repeated with at least four or five individuals, so that measurements, in each case as reliable as this one for 'Addition on Reading Hungarian,' will be available.

His measurements alone, although the product of many hundred hours of experimentation and calculation, do not even permit a reliable measurement of the average effect of all replacements. In general subject A is helped by them so as to fall only half as far below the 'expected ability from unreduced practice' as when the one kind of work is continued. So also the drop below the average performance of the first half hour is, with him, only half as great as when the one kind of work is continued. But, on the average, the other subjects did not benefit appreciably by the changes. The most probably true measurement of the average effect of all the replacements is a loss of 8 per cent compared with a loss of 12 per cent in the case of uniform work. But it is not at all secure.

The reader will, of course, understand that no inference should be drawn from the effect of changing the work of the third and fourth of five 15-minute periods, to the effect of changing work after say two, four, six, or eight hours. An educated adult accustomed to mental work would rarely expect benefit from changing work every half-hour. Nor should inferences from the effect of changing work upon a subject who resolutely does his best throughout be applied to the case of ordinary work in schools and occupations, where ennui and repulsion are the factors which change is expected to reduce.

MEANS OF PREVENTING INJURY FROM OVER-WORK

A certain amount of mental work is healthful. The connection-system requires exercise as truly as food and rest. It can have too little as well as too much activity; and it maintains its 'tone' and power of resistance to mental disease better if a certain amount of its activity is 'work' in the service of remote and unselfish ends, rather than 'play' for personal and immediate gratification.

Too much work may be injurious positively, not only by direct mischief to the neurones doing the work itself, but also by producing in the system the states corresponding to over-excitement and worry. It may also be injurious negatively by depriving the animal of the joy, appetite, physical exercise, and sleep essential to health. It may be injurious in the broader sense of diminishing the value of life, by its deprivations, of whatever sort. As men and things now are, the direct injury intrinsically and necessarily consequent upon mental work, seems to be very, very much less than that due to over-excitement, worry, and the physical, intellectual and moral deprivations.

For over-excitement and worry from mental work, wise formation of habits is the preventive and cure. Mental workers should be taught that emotionality is not a measure of interest, that tension is not a measure of energy, nor over-action of strength, that anxiety is not a measure of devotion, and that peaceful absorption is the feeling proper to achievement. Having learned to judge their efficiency, not by how they feel, but by what they get done, they should practice themselves in casting off every weight of irrelevant thought or feeling, in dismissing as unhealthy and immoral all worry over what has been done, or what one cannot prevent.

For the deprivations, the first remedies to be applied are healthful physical conditions, interest and motive. Proper air and light, proper posture and physical exercise, enough food and sleep, and work whose purpose is rational, whose difficulty is adapted to one's powers, and whose rewards are just, should be tried before recourse to the abandonment of work itself. It is indeed doubtful if sheer rest is the appropriate remedy for a hundredth part of the injuries that result from mental work in our present irrational conduct of it.

However, since for many men, for a long time to come, mental work probably will be carried on with effort against resistance, by individuals who are not properly guarded in general health, it is worth while to inquire whether there is some point or stage in the course of mental work at which a worker should allow himself, or be allowed, or perhaps required, to stop work.

For the unlearned activities and those developing out of them in a simple environment of unconquered nature and of human beings unsophisticated by ideas, there are present certain equally unlearned checks to over-activity. Mental work beyond a certain point produces ennui, repugnance, sleepiness and pain; prolonged restraint from individual or social play produces an intense impulse to its gratification. In the absence of habits of forcing oneself to work in spite of present discomfort and deprivation, these natural checks would operate freely. The animal would be protected against over-work in the same measure that he was protected against starvation or over-feeding, by unlearned impulses. These would work crudely and imperfectly, sometimes failing to check the activity aroused by hunger or the sex instincts, when rest would be preferable, and sometimes letting him rest when continued vigilance would save him.

In the complicated environment created by human intellect and morals, man learns to neglect these natural checks in favor of more remote and civilized ends, and is forced by fear of punishment to work in spite of them. They may fail to operate at all, temporary zeal or long habit rendering the individual immune to all impulses contrary to the accomplishment of his work. It would perhaps be possible for not a single one of these checks to operate, no matter how long work continued, until the man, possessed by zeal for the beloved achievement, and unwarned by repugnance, sleepiness or pain, died cheerfully working to the end.*

* One inference from the mental energy doctrine has been that the mind is protected against overworking by the steady diminution in its efficiency as work continues. The protection thus afforded can be only

It is for the welfare of men in the long run not to obey these *natural* checks, but no one simple *rational* check can be used to replace them. It is consequently impossible to find any uniform rule for deciding when to stop work. 'Follow nature,' 'Work as long as you can,' 'Work until a decrease in efficiency appears,' or any other rule announced for all workers, is bound to be wrong. It is unnecessary for most workers to stop when they are bored and sleepy, and it is unsafe for some to work until they are. The best practical rule seems to be to make sure of adequate exercise and sleep, to divide the balance of time reasonably between the duties and pleasures of life, and to work throughout the amount of time due for work, diminishing the natural checks so far as may be by securing proper physical conditions, interest and motive, and, for the rest, disregarding them. What amount of exercise and what amount of sleep are adequate varies with age and individuality. To insure against injury, the allowance may be made generously. The essence of mental hygiene is then—interest for efficiency; and for protection, sleep.

trivial. The protection afforded by man's original organization, is rather the impluse to stop work outright. This impulse we can and do withstand, with or without feelings of strain and effort. The feelings of strain and effort can in turn be withstood, so that work may continue long after the point at which original human nature would have abandoned it. And when it does so continue, it is maintained at nearly equal efficiency for a long period.

PART II

Individual Differences and Their Causes

PART II

Individual Differences and Their Causes

CHAPTER VII

Introduction

THE PROBLEMS OF INDIVIDUAL DIFFERENCES

In describing the original tendencies of man as a species, attention was called to the fact that the original natures of individual men and women were not exact duplicates, presenting the characteristics of the human species invariably, but deviated from the type of the species in the strength of this, that and the other instinct. In describing the laws of learning or modifiability and the changes in mental functions which learning brings to pass, it was assumed that different individuals learned at different rates; and that identical natures must, if subjected to the action of different external situations or environments, become different. The reports of studies of the amount, rate and permanence of improvement gave frequent illustrations of the variability of individual men in whatever feature of intellect, character or skill we examine. It is the purpose of the remaining chapters of this volume to present the main facts concerning these individual differences and their causes.

Their significance for educational theory and practice is obvious. What we think and what we do about education is certainly influenced by our opinions about such matters as individual differences in children, inborn traits, heredity, sex differences, the specialization of mental abilities, their interrelations, the relation between them and physical endowments, normal mental growth, its periodicities, and the method of action and relative importance of various environmental influences. For instance, schemes for individual instruction and for different rates of promotion are undertaken largely because of certain beliefs concerning the prevalence and amount of differences in mental capacity; the conduct of at least two classes out of every three is determined in great measure by the teachers' faith that mental abilities are so little specialized that improvement in any one of them will help all the rest; manual training is often introduced into schools on the strength of somebody's confidence that skill in movement is intimately connected with efficiency in thinking; the practical action with regard to coeducation has been accompanied, and doubtless influenced, by arguments about the identity or the equality of the minds of men and women; the American public school system rests on a total disregard of hereditary mental differences between the classes and the masses; curricula are planned with some speculation concerning mental development as a guide. It is thus easy to find cases where educational practice depends upon opinions about our group of topics. It is still easier to note a similar dependence in the case of educational theory. Abundant illustrations will appear in the course of our study of the topics themselves.

Exact knowledge of the nature and amount of individual differences in intellect, character and behavior is valuable to educational theory and practice for two reasons. The first is the general need of knowledge of what human beings are in order to choose the best means of changing them for the better —a need which includes knowledge of the divergences of individuals from the type of the species as a whole, as well as knowledge of that type. Education needs knowledge of men as well as of man. The second reason is that by a study of the causes of these differences,—the causes which make men good and bad, wise and foolish, skillful and clumsy, efficient and futile,—education may hope to learn about means of making all men more wise, skillful and efficient. The causes of the differences between one man and another, as things now are, will lead to knowledge of the causes whereby all men may be made to differ from their former selves. It is of special importance to know what differences amongst men are due to differences in sex, race, immediate ancestry and maturity, which are beyond control by ordinary educational endeavors, and what differences, on the other hand, are due to training or education itself.

Effective description of the facts of individual differences and of their causation must be quantitative. The questions are questions of amount, or at least become such when carried beyond a first survey. "Do boys and girls differ?" is itself a question of amount, which soon becomes, "How much do boys and girls differ?" "In what do they differ?" can be answered only by comparing them quantitatively. "Are there distinct types of children with respect to imagination?" can be properly answered only by measuring children in respect to the various sorts of imaginativeness or imagery in question. "What is the value of the study of Latin?" means to even the student most averse to quantitative thinking, "What changes in human nature are caused by it?" But to prove the existence of any change one must measure two conditions.

It is therefore necessary to understand certain elementary facts about the means and methods of measuring the facts of individual differences and their causation, in order to understand the facts themselves. Portions of certain chapters will consequently be given up to the essentials of the theory and practice of measuring mental conditions, differences, changes and relationships.

A CONCRETE ILLUSTRATION OF THE PROBLEMS OF INDIVIDUAL DIFFERENCES

The best means of introduction to the study of individual differences, their causes and their educational significance, will be to examine an actual first-hand study of them. For this purpose I choose certain parts of Mr. S. A. Courtis' report on the arithmetical abilities of children in the schools of New York City ['11-'12].

Mr. Courtis measured the achievements of pupils in responding to eight tests. Test 7 is reproduced below.

ARITHMETIC—Test No. 7. Fundamentals

Name.................*School*.................*Grade*......

In the blank space below, work as many of these examples as possible in the time allowed. Work them in order as numbered, writing each answer in the "answer" column before commencing a new example. Do no work on any other paper.

Number	Operation	Example	Answer	Right	
1	Addition	a 25 + 830 + 122 =......(Write answer in this column) ☞.. b 232 + 8021 + 703 + 3030 = ..			
2	Subtraction	a 5496 — 163 = ... b 943276 — 812102 = ...			
3	Multiplication	2012 × 213 = ...			
4	Division	158664 ÷ 132 = ...			
5	Addition	6134 + 213 + 4800 + 6005 + 3050 + 474 = ...			
6	Subtraction	73210142 — 49676378 = ...			
7, 8	Multiplication	46508 × 456 = ...			
9	Division	27217182 ÷ 6 = ...			
10, 11	Division	3127102 ÷ 463 = ...			
12, 13	Addition	85586 + 69685 + 39397 + 95836 + 37768 + 69666 + 78888 + 54987 = ...			
14	Subtraction	15655431 — 5878675 = ...			
15, 16	Multiplication	78965 × 678 = ...			
17	Division	44502486 ÷ 7 —			
18, 19	Division	5373003 ÷ 769 —			

Consider now the results of Test 7 in a certain eighth-grade class as shown in Fig. 21. Consider also Table 22. which gives similar facts for all the eighth grade children

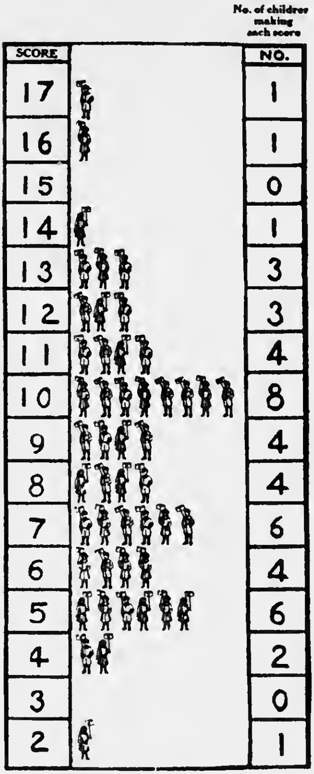


FIG. 21. The Variation in Ability within a Single Class. After Courtis, '11-'12, p. 48.

tested. The picture and this table state an important fact—the existence of great individual differences even among those of the same school grade, and so of roughly similar training in arithmetic—and suggest at least two important questions. The first question concerns the meaning of the score used; the second concerns the comparative effects of (a) differences in original capacities and (b) differences in the circumstances of life and training in producing the differences that are actually found between one human being and another.

TABLE 22.

THE VARIATION AMONG 8TH GRADE PUPILS IN ARITHMETICAL COMPUTATION. After Courtis, '11-'12, p. 46.

"Score" or "Quantity": Examples done correctly in 12 minutes in the case of Test 7	Number of children making each score or "Frequency": 8th grade children in New York City
19	31
18	25
17	86
16	107
15	182
14	251
13	327
12	390
11	453
10	497
9	475
8	425
7	333
6	312
5	239
4	152
3	88
2	71
1	30
0	28

The score used here is the number of examples done correctly (certain more difficult examples counting each as 2). Elsewhere Mr. Courtis reports also the number of examples "attempted" but not completed with a correct answer. Now, obviously, the individual abilities reported in Fig. 21 and in Table 22 might be more widely scattered or be more closely clustered if some credit were given for the amount achieved in each example not completed with a correct answer. Moreover, obviously, 'to do correctly 16 examples' need not mean to have eight times as much ability, or to achieve eight times as much, as 'to do correctly two examples in the same length of time,' in any absolute sense. As I have already repeatedly suggested in the previous volume, we cannot assume that, if A takes eight times as long as B to run a mile, to write an equally good epic poem, to discover the cure for cancer, to translate a page of Latin, to add a column of figures, to eat his meals, and to make up his mind to give all that he has to the poor, B's achievements or abilities as runner, poet, scientist, translator, adder, and B's appetite and charity, are in each case eight times as far above zero as A's.

We have seen that where an individual's condition at one time is compared with his condition at another, the comparison must be made with recognition of the nature of the score used. So, also, when one individual's condition is compared

TABLE 23.

THE OVERLAPPING OF LOWER AND HIGHER GRADES IN ARITHMETICAL COMPUTATION. After Courtis, '11-'12, p. 46.

Score or Quantity:	Frequency		
Examples done correctly in 12 minutes in the case of Test 7.	In the 4th grade	In the 6th grade	In the 8th grade
19	1	3	31
18			25
17	1	10	86
16	1	17	107
15		30	182
14		62	251
13	6	106	327
12	9	134	390
11	31	304	453
10	42	358	497
9	132	505	475
8	280	647	425
7	405	685	333
6	487	620	312
5	645	571	239
4	661	494	152
3	690	388	88
2	654	306	71
1	628	203	30
9	723	227	28

with another's. All thought about individual differences must allow for the meaning of the scales and scores used.

The second problem suggested—of the division of responsibility for individual differences between original nature and environment or nurture—can be only suggested, not solved, by these data themselves. The wide range of achievement of pupils of roughly the same school training suggests that differences born in them play a large part in determining the differences eventually found in them. This is emphasized by the extensive 'overlapping' of the scores from pupils of any grade, by the scores from pupils of lower grades shown in Table 23. Thus, some of the pupils in Grade 4 made even better scores (in number of correct answers obtained) than the average pupil of Grade 8. The difference between the best and the worst score in the same grade is three times the difference between the average score of the eighth grade and the average score of the fourth grade. And we shall later find other evidence leading to the same conclusion. The problem of the exact relative shares of nature and nurture is, however, a very complex and difficult problem, requiring elaborate consideration in later chapters.

Another problem in the causation of individual differences is illustrated by Mr. Courtis' tables comparing the two sexes. I quote (in Table 24) the one for Test 6 in the 7B grade.*

It appears that in the number of examples attempted there was little or no difference between the sexes, but that in the number of correct answers the boys did somewhat better than the girls of the same grade.

The effects of sex, whether by inherited sex qualities or by the circumstances in which training differs for the sexes, have been the subject of many speculative opinions and of some few

* Test 6 was as follows:

Do not work the following examples. Read each example through, make up your mind what operation you would use if you were going to work it, then write the name of the operation selected in the blank space after the example. Use the following abbreviations:—"Add." for

Table 24.

Sex Differences in a Speed Test in Reasoning: Boys and Girls in the 7 B Grades Compared. After Courtis, '11-'12, p. 138.

Quantity: Number of Examples attempted in Test 6 in 1 minute	Frequency in 7 B Grade		Quantity: Number of Examples done correctly in Test 6 in 1 minute	Frequency in 7 B Grade	
	Boys	Girls		Boys	Girls
16	1	6	16	1	
15			15		
14	1		14		
13	4	3	13		
12	7	6	12	1	1
11	7	11	11		2
10	15	18	10	6	4
9	34	21	9	10	4
8	59	52	8	23	11
7	119	88	7	50	24
6	240	216	6	110	61
5	287	273	5	197	132
4	238	230	4	245	197
3	161	172	3	245	236
2	55	65	2	201	273
1	6	6	1	113	175
0	1	1	0	33	48

addition, "Sub." for subtraction, "Mul." for multiplication, and "Div." for division.

	Operation	
1. The children of a school gave a sleighride party. There were 9 sleighs used, and each sleigh held 30 children. How many children were there in the party		
2. Two school-girls played a number game. The score of the girl that lost was 57 points and she was beaten by 16 points. What was the score of the girl that won?		
3. A girl counted the automobiles that passed a school. The total was 60 in two hours. If the girl saw 27 pass the first hour how many did she see the second?		
4. On a playground there were five equal groups of children each playing a different game. If there were 75 children altogether, how many were there in each group?		

And so on for twelve similar examples.

impartial investigations. They will be discussed by themselves in Chapter IX.

The influence of remote ancestry or race could be studied similarly by comparing two groups of children of the same sex, age, and training, but different in that the one group sprang from, say, East European Hebrews and the other from, say, North American Indians.

Other possible causes of differences in arithmetical achievement are:—differences in near ancestry or 'family,' differences in maturity, and differences in the length of time devoted to arithmetic in the schools, in the methods used in teaching it, or in other circumstances of training.

CHAPTER VIII

The Measurement of Individual Differences

SIMPLE AND COMPOUND DIFFERENCES

A difference in human nature may be (1) in the amount or degree of the same thing* (as 'good—better,' 'quick—slow,' 'imaginative—less imaginative—unimaginative'); or (2) in the presence or absence of different things (as 'John knows Latin; James knows German,' or, 'A is imaginative; B is rational,' or, 'C has an artistic temperament; D has a scientific temperament').

The second case commonly reduces to an aggregate of differences of the first sort if the statement of difference is made adequate. Thus, 'John knows *x* Latin; James knows *o* Latin; James knows *y* German; John knows *o* German' should properly replace the former statement. Similarly we have, 'A is imaginative to *z* extent, B is imaginative to *z-w* extent. B is rational to *v* extent; A is rational to *v-w* extent,' and, 'C has *r*, *y* and *z* amounts, respectively, of certain qualities, certain degrees of which in combination we call the artistic temperament; D has small or possibly zero amounts of these qualities.' A difference in human nature then commonly is a difference in the amount of one thing or an aggregate of differences each in the amount of one thing.

But conceivably there may be things which do not vary in amount except from *zero* by a sudden jump to one positive condition. The thing would then either be in one constant degree or not at all. The literature of psychology and of edu-

* Or what is assumed to be the same thing.

cation abounds in cases of difference stated as if the difference were that between *o* and *k*, without k-1/10, or k-2/10, or o+1/8k, etc. But such statements are usually due to ignorance or vagueness. 'John is color blind; James is not' cannot really mean that there is some one constant degree of color blindness which a man either has just in that degree or does not have at all; for there are varying amounts of color blindness. In fact, it is doubtful if there are any individual differences in human intellect or character of this '*zero* to *k*' sort.

When, therefore, it is stated that John and James differ in *kind,* the statement always, or almost always, means nothing more than that one of the two individuals possesses a certain amount or degree of something of which the other possesses *o*. When it is stated that the difference between John and James in one thing, say, knowledge, is *qualitative* whereas their difference in another, say, motor skill, is *quantitative,* the statement always, or almost always, means nothing more than that certain sorts of knowledge, present in certain amounts in John, are absent in James, and vice versa; whereas, in motor skill, both have perceptible amounts of an identical set of things.

It is then not only permissible, but more scientific and more useful, to think of human individuals as all measured upon the same series of scales, each scale being for the amount of some one thing, there being scales for every thing in human nature, and each person being recorded as *zero* in the case of things not appearing in his nature. And the only problem of method which need concern us in this chapter is the problem of the nature and use of a scale for measuring different amounts or degrees of the same trait in different human beings.

UNITS AND SCALES FOR MEASURING MENTAL DIFFERENCES

The facts of importance about scales for mental and moral traits can be best stated in connection with some concrete illustrations.

Consider the facts given in Table 25 concerning individuals A, B and C, with reference to these questions:—

1. What are the differences between A and B, A and C, and B and C in each of the traits?
2. How many times as great is the difference between A and B in trait I as that between B and C in trait I? How many times as great is the difference between B and C in trait II as that between A and C in trait II, etc., etc.

What the difference is between A and B can be definitely determined in the case of stature (I), reaction-time (II), error in drawing a line (III), and age (XII); but all that one learns about the differences in the case of ability in history (VII) and interest in music (XI) is that one difference is that between *excellent* and *good* on the arbitrary scale of some school and that the other is that between *little* and *moderate* in the mind of some observer. The measures of I, II, III and XII are by *objective* or *impersonal* scales, that is, scales the identification and similar use of which are possible for any competent observer. The measures of VII and XI are by *subjective* or *personal* scales, which another observer could not identify* or use in the way in which the person giving the marks used them. Moreover, whatever the scales in VII and XI really are, they are certainly very coarse; a wide range of difference is expressed in them by a few steps or marks or values. In short, any competent thinker knows exactly what is meant by 160 cm. and by 160 cm.—140 cm., whereas he cannot be sure what is meant by *excellent,* or by the difference between *excellent* and *good.*

Concerning the other measurements the following statements are roughly true:—

A, B and C are measured objectively in the quality of handwriting. We know what is meant by the difference between

*If, for example, the reader had heard the oral work and seen the written work which in combination mean *good* in the school whence A and B came, he would not know that they did mean *good.*

TABLE 25.

MEASUREMENTS OF THREE INDIVIDUALS, A, B, AND C.

	A	B	C
I. Stature	160 cm.	140 cm.	130 cm.
II. Simple reaction-time to sound	.175 sec.	.125 sec.	.150 sec.
III. Average error in drawing a line to equal a 100 mm. line	3.2 mm.	2.8 mm.	2.2 mm.
IV. Number of words (of a list of 12, heard at a rate of 1 per second) remembered long enough to write them immediately after the last word was read	6 words	9 words	7 words
V. Number of examples in addition (each of 10 numbers, repeating no number in any one example, taken at random from the numbers 10 to 99) done correctly in 8 minutes	14	12	18
VI. Quality, or merit, or goodness of handwriting	See Fig. 22	See Fig. 23	See Fig. 24
VII. School marks in history	Ex.	Good	Poor
VIII. School marks in spelling	82	62	93
IX. Efficiency in perception; the number of A's marked in 60 seconds on a sheet containing 100 A's mixed with 400 other capital letters	48 A's	60 A's	82 A's
X. Criminality: number of times convicted of a penal offense	0	1	0
XI. Degree of interest in music	little	moderate	a great deal
XII. Age in days	5080 d.	6150 d.	5615 d.

A and B. It is just the difference between the qualities of the two samples presented. But they are not measured conveniently; the differences are not referred to any commonly known scale.

The measurement in addition is nearly objective. If the conditions of the test are defined by a statement of how the examples were presented (what type they were printed in, how they were arranged, etc.), how the answers were given, at what time, under what distractions, with what incentives, etc., they were done, and the like, they become still more fully objective.

In the case of memory, not only the conditions of the test, as in addition, but also the exact words used need to be specified.

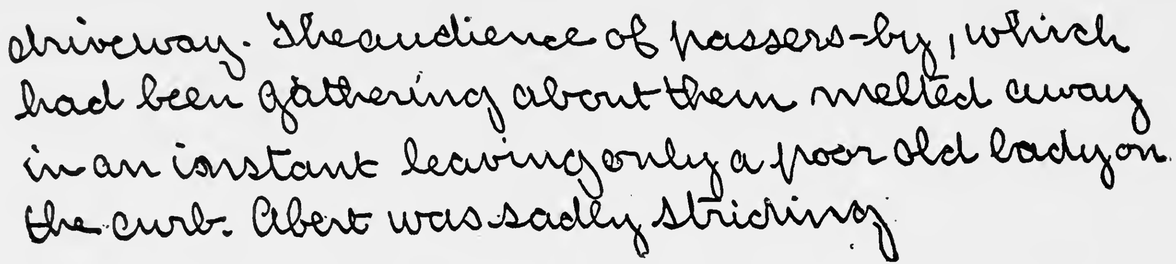


Fig. 22. A's Handwriting.

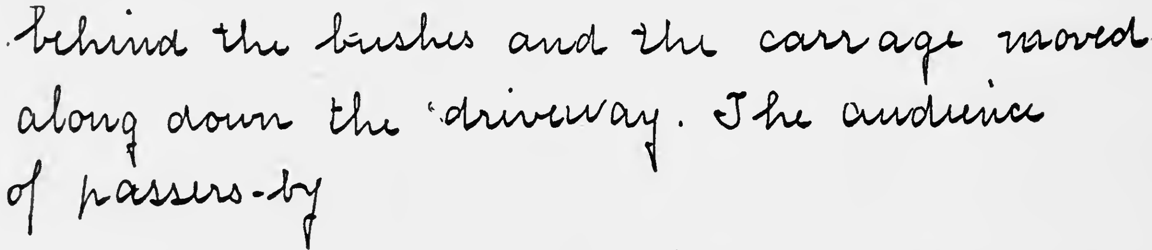


Fig. 23. B's Handwriting.

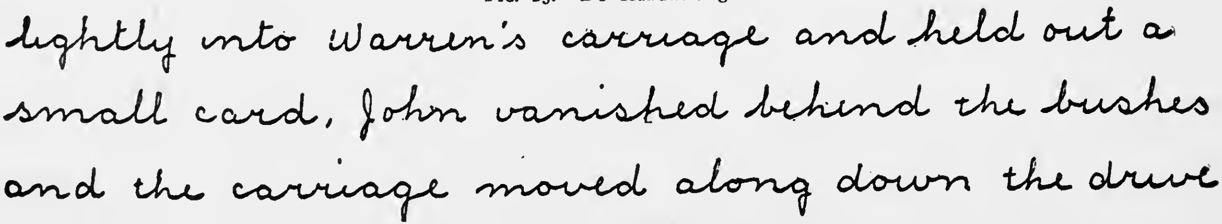


Fig. 24. C's Handwriting.

A differed from B by writing down three less words of a certain list. The list needs to be known if all competent observers are to think of the same thing by *three words.*

The measure of criminality is inferior to the measurement of stature or age by using an ambiguous unit (*convicted of a penal offense*) and also by measuring only very coarse differences. A and C are not differentiated by the measures, though A might just fall short of crime and C be a very healthy-minded and kind-hearted boy.

The measurement of ability in spelling might turn out to be, were the conditions of the test and the system of scoring results in it known, as objective as that in addition, or, on the other hand, may be only a record of the opinion of some teacher that A was a good deal better than B and that C was somewhat above A. As the record stands the $A - B = 20$ may mean no more than '*A is good; B is unsatisfactory,*' or '*A is somewhat better than B in my opinion.*'

The values $\frac{A - B}{A - C}$, $\frac{B - C}{A - C}$ and $\frac{A - B}{B - C}$ are determinable in the case of I, II, III and XII of Table I. In stature $A - B = 20$ cm., $A - C = 30$ cm. and $\frac{A - B}{A - C} = .667$. When one says that A differs from C in stature one and a half times as much as from B, any competent person knows exactly what is meant. For VII on the contrary we have $\frac{A - B}{A - C} = \frac{\text{Excellent} - \text{Good}}{\text{Excellent} - \text{Poor}}$, which may be a true, but remains a mystical, answer, until *excellent, good* and *poor* are defined on some scale.

In the case of the memory of words ($A = 6$, $B = 9$, $C - 7$) $\frac{B - A}{C - A} = 3$ and $\frac{B - C}{C - A} = 2$, if *to remember any one word* = *to remember any other one word.* Suppose, however, that the series of words was: '*career, dilatory, opium, never, soap, numbers, add, subtract, one, two, three, four,*' and that A remembered the last 6, C the last 7, and B these and the first two. It would be very risky to assume that the difference between A and B (remembering *career, dilatory* and *numbers*) was only three times the difference between A and C (remembering *numbers*). And it would be almost certain that a difference of 10—4 in the test would be really more than twice as great as a difference of 7—4. Under the circumstances *numbers, add* and *subtract* are probably much easier to remember than *career, dilatory* and *opium.*

In the case of centimeters or seconds every competent person knows not only what fact is meant by any given number of the units, but also that *any one unit is equal to any other one, any two to any other two, and so on.* In the case of words remembered, units called by the same name may not be really equal.

In the case of the addition, one example differed from another in difficulty only by chance and only slightly; for when 10 two-place numbers are picked by chance and arranged in a chance order, the chances are enormously against getting any very easy examples like (a) or any very hard examples like (b).

(a)	(b)
25	89
35	95
30	58
40	67
60	79
—	—

The chances are still more against getting two or three examples in succession that are on the average more than a trifle harder than any other succession of two or three. Reliance upon the truth of $\frac{A - B}{C - B} = \frac{2}{6}$ would, however, be safer if 10, 20, 30, etc., and 11, 21, 31, etc., had been excluded from the numbers used in the tests.

In the case of merit of handwriting, the differences between A and B, B and C, and A and C do not even pretend to be put in terms so as to allow comparison. As the records stand, any one must get the differences transposed into terms of some unit before he can calculate $\frac{A - B}{A - C}$ at all. Since the measures of A, B and C are objective, this can be done. If, for example, the measurer could show that the difference between A and C was approximately five-twelfths of the difference between two standard samples accessible to all competent persons, and that the difference between A and B was approximately two-twelfths of the difference between the same two standard samples, he could then regard $\frac{A - B}{A - C}$ as, of course, $\frac{-2/12\,k}{-5/12\,k}$ or $\frac{2}{5}$, k being the difference between the two standard samples. As a matter of fact the difference between the sample of Fig. 22 and that of Fig. 24 is in the combined opinion of some hundred judges just about two and one-half times as great as the difference between the sample of Fig. 22 and that of Fig. 23.

The sum and substance of the last four pages is that a measurement of human nature to be useful for our purpose must identify the *amount* in question for any competent thinker, just as a useful description must *identify* the *object* in question. To do this it must be objective, that is, free from individual caprice, so that any competent person making the same measurement would get the same result.*

* Not, of course, exactly the same. There is a personal equation in even the most objective measures, such as the length of this line ————. If they measured it to thousandths of a millimeter, competent observers would not get the same result, except by chance. Nor would the same observer in several independent measurements. The ultimate distinction between objective and subjective is simply that in the former sort of measurements competent observers use very nearly the same criteria and, tho independent, agree very closely, whereas in the latter they use very different criteria and, if independent, agree only roughly. The reader may safely postpone any subtle or thoroughgoing treatment of the distinction until he can study the theory of mental measurements in detail. For the purposes of this book, objective measures may be defined as measures which competent observers could repeat and verify or reject, and subjective measures as measures which they could not so repeat and verify.

It should also, if possible, so state the amount in question that other amounts of the same thing may be compared with it as so much greater or less, permitting differences between the amounts to be expressed in ratios. In still briefer terms, the measurements should be *at defined points on an objective scale, the distances of these points one from another being also defined.*

The reason for the elaborate introduction to and illustrations of this obvious principle is that in its application to particular problems in the study of educational psychology it has not been obvious to even the writers of treatises and investigators of original data, much less to the rank and file of students of psychology or of education. On the contrary this entire chapter would not suffice to list and barely describe the quantitative conclusions that have been drawn from subjective opinions or from the unjustifiable acceptance as equal of units which happened to be called by the same name.

The reason for contrasting physical and mental measurements, to the apparent disrepute of the latter, is not that I wish to discourage the reader from trusting, or from making, measurements of any feature whatever of intellect or character. It is to be hoped, however, that he will be effectually discouraged from trusting measurements which do not deserve trust. The lesson to be drawn from the contrast is that a measurement can rightly be trusted or rejected or criticized or, indeed, understood, only if the concrete reality which it describes is known. Any numerical statement has meaning only in reference to a concrete scale and its units.

THE VARIABILITY OF A MENTAL MEASUREMENT

One cause of improper distrust of measurements of intellect and character is so common that it demands special treatment. This is the variability of the same measurement of apparently the same fact. For instance, individual A was tested by hearing a series of 12 letters read at a rate of 2 per second, he being required to write down as many as he could remember in their proper order as soon as the reading was finished. His score was 4 correct in one trial and 10 correct in another, the two letter-series being, for people in general, of equal difficulty. What assurance, it may be said, can be felt in a measurement of now 4 and now 10 for the same fact? The defense is that it is not the same fact. To measure A's memory for series of letters is not to measure one constant thing, but a very variable thing. A, as regards taking in and holding a series of letters, is *not* the same from moment to moment. What is measured in one trial is a sample of A's varying status in respect to this ability. His first score in the measurement in question was 4. This measurement is, so far, so good; it is better to believe that A's ability in that test is 4 than to guess at it. In a second test the score was 10.

To say that A's ability is 4 or 10 or averages 7 is better than to have taken the 4 as a measure. But further trials give (including these two)

1	record	of	4	letters	correct
4	records	"	5	"	"
4	"	"	6	"	"
7	"	"	7	"	"
13	"	"	8	"	"
3	"	"	9	"	"
4	"	"	10	"	"

From all these scores we get, as an average, 7.44 letters correctly written; as the most common record (the so-called Mode), 8 letters correctly written. Since a record of 4 letters correctly written means, perhaps, 4½ letters remembered, one of 5 letters correctly written, 5½ letters remembered and so on, we should perhaps call A's average memory for letters 7.94, and his mode or most frequent performance 8.5.*

If an individual writes 7 letters correctly, but does not write an eighth, it means that he remembered at least 7 and not 8. The measurement is comparable to one of length in which the observer notes that a stick is over 62 inches, and is not 63 inches, long. And just as, in the long run, sticks 62 inches or over but not 63 inches or over will average 62½ inches, Now the trustworthiness of any one of these is 6 times as great as that of the first single score 4. But we can not be sure that the average of the 36 measurements is identical with A's true average ability. In fact we can be almost sure that it is not. Seventy-two measurements might give, and almost certainly would give, a slightly different average. True average ability in the case of variable measurements means the measure we would get as an average from an infinite number of measurements. Only by chance will the result from any finite number of measurements be identical with it. All our measures represent approximations, but the greater the number of measures the closer the approximation will be.

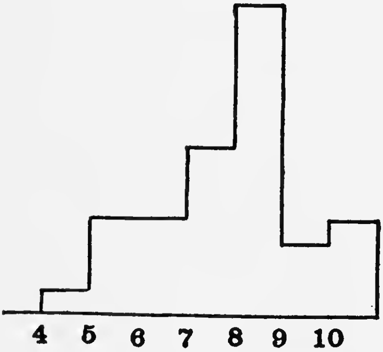


Fig. 25. The Ability of Individual A in Memory of Letters.

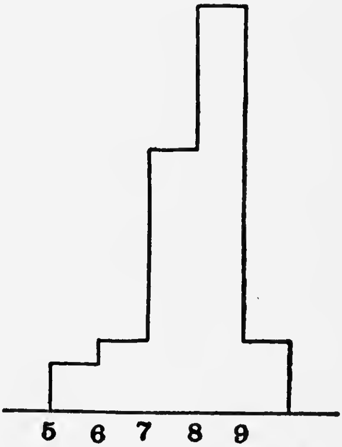


Fig. 26. The Ability of Individual B in Memory of Letters.

The variability of mental measurements thus gives no reason to distrust them, but, on the contrary, gives a means of knowing just how trustworthy they are.

so the person remembering on many occasions at least 7 but not 8 letters will average 7½ letters. Some of his records of 7 mean 7 letters remembered and one more letter almost, but not quite, remembered.

* It is found in variable measurements of the ordinary sort that the reliability of an average increases as the square root of the number of measurements taken.

The series of measures above is our knowledge of A's ability. We can see the fact more clearly by expressing it in space rather than in figures. If we let each quarter inch along a horizontal line stand for one letter correctly written, and each eighth of an inch of height above it stand for one manifestation by A of the ability designated by that place, we have Fig. 25, by which one can see at a glance A's ability, its variability and his general tendency to keep nearer 8 than any other one ability.

If we must for any reason abbreviate our description of A's ability we may best take two measures, one of the ability about which his various scores center most closely and the other of the closeness of his grouping. We may term these the *central tendency* and the *variability*. For the former the *average* or *median** or *mode* may be used; for the latter, the average of the differences between the individual records and their central tendency (Average Deviation or A.D.), or any one of a number of measures of the closeness of clustering of the individual records about their central tendency.

Let us suppose that, with the same test, individual B showed the following ability:

2	records of	5	correct.
3	"	6	"
11	"	7	"
17	"	8	"
3	"	9	"

This is shown graphically in Fig. 26. The average, median and mode would be closely the same as for A, but the varia-

*The Median has two meanings: the point on the scale above and below which equal per cents of the individual scores lie, and the mid measure, that is, the $\frac{n+1}{2}$ th measure, counting in from either extreme, where $n =$ the total number of measures. These two definitions lead to substantially the same results, and for present purposes the reader may adopt either one. Indeed it will do no harm if he can see no difference between the two.

bility of the measures would be less. The limits for A were 4—10. For B they are 5—9. The average difference of the individual measure from the mode for A was 1.17. For B it is 0.7. B's ability has the same central tendency as A's, but B is a more constant performer.

TABLES OF FREQUENCY OR DISTRIBUTION

Suppose now that the average ability in remembering letters in such a test as that described had been determined for every human being six years old or older. From such records the difference of any individual from any other could be computed, but only by hunting out the records of the two individuals. The frequency of any given *degree* of difference could be found from a very simple summary of such records, such as appears in Table 26.

TABLE 26.

SUPPOSED DISTRIBUTION OF AVERAGE ABILITY TO REMEMBER LETTERS IN THE CASE OF HUMAN INDIVIDUALS SIX YEARS OLD OR OLDER.

An average of			0 letters	was remembered by			0	individuals
"	"	"	1 letter	"	"	"	2,000,000	"
"	"	"	2 letters	"	"	"	30,000,000	"
"	"	"	3 "	"	"	"	60,000,000	"
"	"	"	4 "	"	"	"	190,000,000	"
"	"	"	5 "	"	"	"	280,000,000	"
"	"	"	6 "	"	"	"	360,000,000	"
"	"	"	7 "	"	"	"	310,000,000	"
"	"	"	8 "	"	"	"	190,000,000	"
"	"	"	9 "	"	"	"	40,000,000	"
"	"	"	10 "	"	"	"	6,000,000	"
"	"	"	11 "	"	"	"	400,000	"
"	"	"	12 "	"	"	"	100,000	"
"	"	"	13 "	"	"	"	40,000	"
"	"	"	14 "	"	"	"	20,000	"
"	"	"	15 "	"	"	"	4,000	"
"	"	"	16 "	"	"	"	500	"
"	"	"	17 "	"	"	"	100	"
"	"	"	18 "	"	"	"	5	"
"	"	"	19 "	"	"	"	0	"
"	"	"	20 "	"	"	"	0	"

A difference of 17 occurs only 10,000,000 times, between each of the 5 individuals of ability 18 and each of the 2,000,000 of ability 1. A difference of 16 occurs 350,000,000 times (5 times 30,000,000+100 times 2,000,000). A difference of 15 occurs 4,300,000,000 times (5 times 60,000,000+100 times 30,000,000+500 times 2,000,000).

A distribution or table of the frequencies of the *abilities* of individuals is then a convenient means of presenting the facts from which the frequencies of the different *differences* amongst

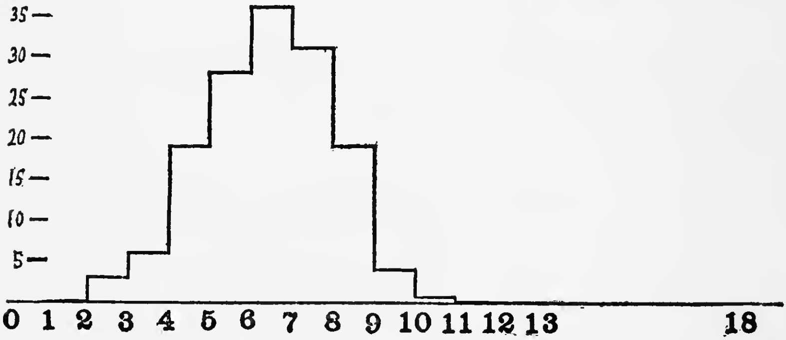


FIG. 27. Supposed Relative Frequencies of the Different Abilities in Remembering Letters of Human Individuals Six Years Old or Older. The horizontal scale is for the average number of words remembered. The vertical scale is for the number of individuals, 10,000,000 being the unit.

them may be calculated. It serves many other purposes as well, and so is commonly used in reporting the results of the measurement of any one trait in a number of individuals. The main features of such a table can be seen at once in their relations one to another if it is presented in graphic form. Thus Table 26 becomes Fig. 27, by letting the amounts of the trait be represented along a horizontal line and the number of persons possessing each amount be represented by the heights of a column erected at the place on the scale denoting that amount. Such a graphic representation is called a *surface of frequency* or a *surface of distribution;* the line which, with the base—or scale—line, encloses this surface is called a *frequency curve* or *distribution curve.*

Real distribution tables and the corresponding surfaces of frequency for certain specified groups of individuals are shown in Table 27 and Figs. 28, 29 and 30 for the cases of *Grade reached by Connecticut children at the age of 10, Speed of sixth-grade children in copying figures,* and *Efficiency of high-school children in adding pairs of one-place numbers.*

TABLE 27.

SAMPLE OF DISTRIBUTION TABLES.

Distribution of Progress in School among Connecticut children in 1903		Distribution of Ability in Copying Figures in 6th grade children. After Courtis, '11-'12, p. 54		Distribution of Ability in Adding Pairs of One-Place Numbers in High-School Pupils. After Courtis, '11-12, p. 52	
Quantity: Grade reached at age of 10	Frequency: Number of Children	Quantity: Number of digits copied in 60 seconds	Frequency in 6th Grade Children	Quantity: Number of pairs added in 60 seconds	Frequency in High School Pupils
		0 to 9	9		
Kinder-		10 " 19	12		
garten	9	20 " 29	22	20 to 29	2
1st grade	442	30 " 39	18	30 " 39	4
2nd "	1389	40 " 49	57	40 " 49	41
3rd "	3293	50 " 59	107	50 " 59	113
4th "	4433	60 " 69	291	60 " 69	272
5th "	3200	70 " 79	536	70 " 79	235
6th "	1227	80 " 89	1274	80 " 89	196
7th "	237	90 " 99	1256	90 " 99	86
8th "	48	100 " 109	1066	100 " 109	43
9th "	4	110 " 119	494	110 " 119	2
10th "	1	120 " 129	359	120 " 129	2
		130 " 139	64		
		140 " 149	36		
		150 " 159	19		
		160 " 169	47		
		170 " 179	2		
		180 " 189	1		

Such tables of distribution or surfaces of frequency are the terms in which the student of individual psychology must do very much of his thinking, both when he tries to describe, and when he tries to account for, the variation of human beings around the type of the species. In the next chapter, for example, in comparing the sexes, in respect to one or another

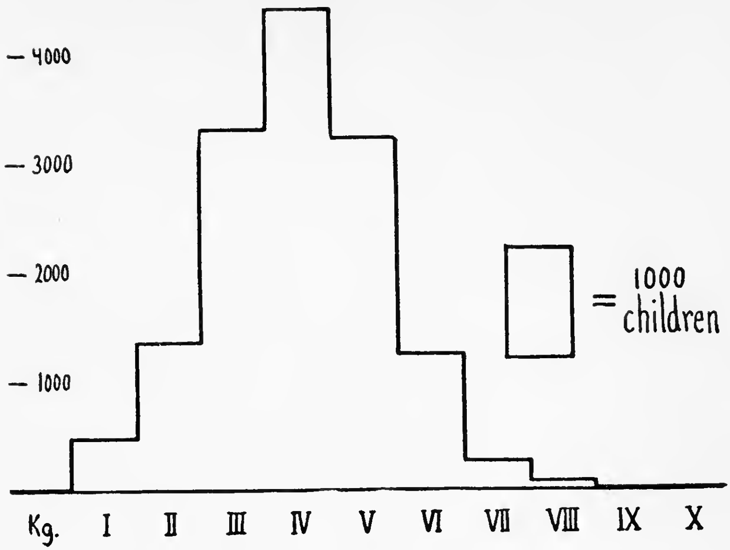


FIG. 28. The Number of 10-year-old Children in Connecticut (1903), in Each School Grade.

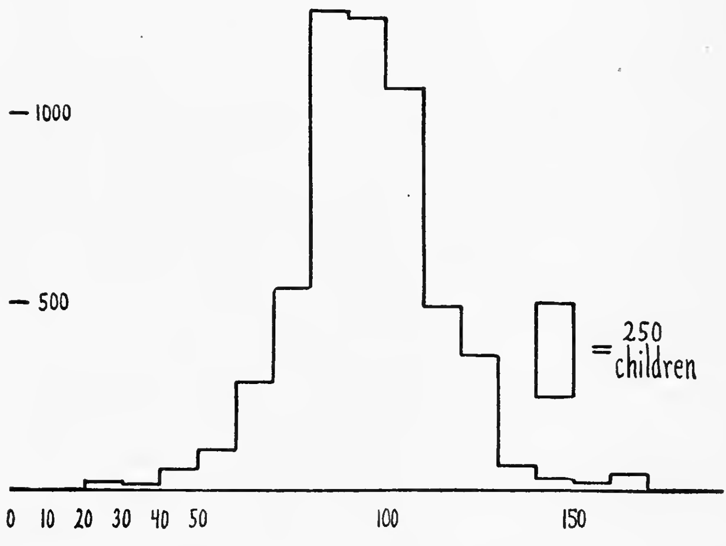


FIG. 29. The Number of Sixth-Grade Children in New York City Copying 0-9 Digits, 10 to 19 Digits, etc., in 60 Seconds.

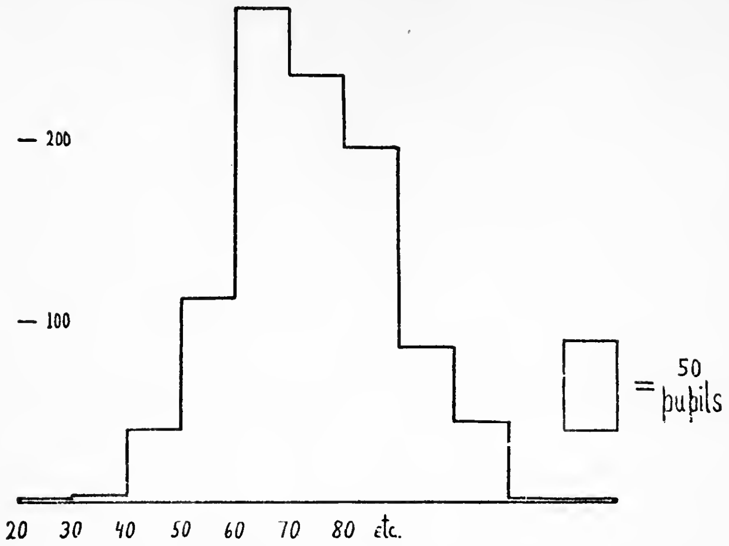


FIG. 30. The Number of High-School Pupils Adding 20-29 Pairs of Numbers, 30-39 Pairs, etc., in 60 Seconds.

mental trait, we shall have to think of the surface of frequency for men in each trait and the surface of frequency for women in the same trait and to compare these two total surfaces of frequency.

CHAPTER IX

The Influence of Sex

SEX DIFFERENCES IN ABILITY

By way of preface to an account of sex differences it is well to note that their existence does not necessarily imply in any case the advisability of differences in school and home training, and, on the other hand, that even if the mental make-up of the sexes were identical it might still be wisest to educate them differently. It is true that a difference of two groups in a mental trait will theoretically involve differences in treatment, but practical considerations apart from that of developing the highest efficiency in that trait may outweigh the advantages of the differential treatment. For instance, consumptives theoretically need a different mode of life from people with healthy lungs, but it might in some cases be wiser to leave a consumptive to his ordinary habits rather than to cause in him consciousness of his disease and worry concerning it. On the other hand, two boys might be identical in mental structure, yet their education might best be very different if we wished to make one of them a chemist and the other a psychologist.

Let us note in the second place that the existence of differences need not imply the need of different training, because those very differences may have been due to the different training actually received and might never have appeared had training been alike in the two classes. It is folly to argue from any mental condition in an individual or class without ascertaining whether it is due to original nature or to training.

The chapter should properly be devoted exclusively to the differences necessarily produced by sex. Those produced by virtue of the adventitiously different training which boy and girl undergo belong in Chapter XIII. So far as may be, such a separation of differences due to sex-nature from those due to our traditional treatment of the sexes is in fact made. But in many cases where the amount of the difference that is to be credited to training is doubtful, the difference will be described in the present chapter, the discount to be made being left to the reader's judgment.

A further caution is necessary before this description and incomplete analysis begins. It is not to confuse differences in behavior, achievement and mental activities indirectly caused by physical traits with such differences directly caused by mental traits. Lack of muscular strength and the phenomena intimately associated with bearing children may serve as samples of such physical traits. Even if women possessed mental capacities for business identical with those of men, they still might not in active work do as much.

In the fourth place, the fallacy of unfair selection must not be forgotten in our comparisons of men and women. For instance, any inference from a comparison of young men and women in college or of working women with men in the same profession is untrustworthy. College women and college men are two classes selected by different agencies. The intellectual impulse has been relatively a more powerful agent in sending girls to college, while convention and the demand for a pleasant social and athletic life have acted more powerfully on boys. In the case of an industry, say laundering, women are selected by relative ignorance, strength, widowhood, drunken husbands, etc., while the men are selected largely by Chinese birth. Let not the bizarre nature of this particular illustration blind us to the fact that women and men physicians, lawyers, stenographers, teachers or government clerks represent different samplings of the two sexes. It is possible theoretically to make a discount for the differential influence of selective agencies and thus permit a fair comparison, but the amount of the discount is very hard to determine.

The investigator of the direct share of sex in the production of mental differences would like to compare individuals alike in age, race, immediate ancestry and training, different in sex alone. The nearest approach that he can make to this crucial comparison is to compare a brother with his twin sister in the case of families where the treatment of the two is most alike.* What he has done is to take such measurements as he can get, of boys and girls or men and women as nearly alike in age and race and training as practical exigencies allow, and to measure enough individuals to make the average immediate ancestry of the one group nearly like that of the other.

A Sample Study of the Influence of Sex

I shall report as a sample of such studies, that of Dr. Thompson. [Thompson, H. B., '03] This report will be followed by a statement of present knowledge concerning sex differences, first in intellectual or semi-intellectual abilities, including sensory and motor abilities, and then in those interests, tendencies and propensities which constitute what we roughly call character and temperament.

Dr. Thompson describes the essential arrangements made by her to secure a just measurement of sex differences, as follows:—['03; pp. 2-6, *passim*]

"In order to make a trustworthy investigation of the variations due to sex alone, therefore, it is essential to secure as material for experimentation, individuals of both sexes who are near the same age, who have the same social status, and who have been subjected to like training and social surroundings.

* I have made this comparison in the case of ten pairs of twins from 9 to 15 years old, in simple but fairly precise tests of efficiency in perceiving small details (A test, a-t, r-e, and misspelled word tests) and of efficiency in controlled associations of ideas (opposites test, addition and multiplication). The difference between boy and girl of a pair of twins varies greatly, but the general result is an absence of difference, the boys doing worse by 1 per cent in the former and better by 2 per cent in the latter tests. This general result might change if ten thousand instead of ten pairs were studied, but the chances are over 9 out of 10 that there would not be a difference of 15 per cent in favor of either sex.

The complete fulfilment of these conditions, even in the most democratic community, is impossible. . . . Probably the nearest approach among adults to the ideal requirement is afforded by the undergraduate students of a coeducational university. For most of them the obtaining of an education has been the one serious business of life. They have had at least the similarity of training and surroundings incident to school life. Most of those in a western university have received their education in coeducational schools.

The individuals who furnished the basis for the present study were students of the University of Chicago. They were all juniors, seniors, or students in the first year of their graduate work. The original intention was to limit the ages to the period from twenty to twenty-five years. Owing to the difficulty of obtaining a sufficient number of subjects within these limits, a few individuals of nineteen years, and a few over twenty-five were admitted. The subjects were obtained by requesting members of the classes in introductory psychology and ethics to serve. They were told nothing about the object of the tests except that they were for the purpose of determining psychological norms. The series of questions on age, health and nationality, ... shows that in all these respects the men and women tested were closely comparable. . . . .

The series of tests employed in this investigation required from fifteen to twenty hours of time from each subject. The hours were arranged from one sitting to the next according to the convenience of the subject. It was not possible to have the hours for any one test constant for all subjects, since the schedules varied so widely. No attempt was made to keep the order of experiments rigidly the same for all. Convenience and economy of time necessarily determined the order to a great extent. In general, however, the simple sensory and motor tests were given in the early part of the series, and the intellectual tests in the latter part. The questions on personality usually came last. The taste and smell experiments had to be scattered through most of the periods, since only a few at a time could be performed without fatigue. The entire series was applied to fifty subjects, twenty-five men and twenty-five women.

The experiments fell into seven groups, dealing respectively with motor ability, skin and muscle senses, taste and smell, hearing, vision, intellectual faculties, and affective processes. . . .

A few words in general on the methods employed may not be out of place, in spite of the fact that each is described in full in connection with the test. The guiding principle in selecting the method was the desire to make the directions to the subject as clear and simple as possible and at the same time secure the greatest possible accuracy of result."

The following quotations give an idea of details of method in the case of two of the tests of "intellectual faculties";— —[Thompson, '03, pp. 111-114]

Test I for Ingenuity

"Fifteen matches were laid on the table in such a way that they formed five squares in the relative position shown in

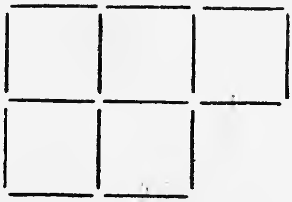


FIG. 31.

Fig. 31. The subject was then asked if he had ever seen the figure before or knew its purpose. One of the fifty—a woman —had seen it before, but had forgotten its purpose. She found the solution in ten seconds, but since she was doubtless assisted by her previous acquaintance with the figure, her record is not included in the curve. The others, upon stating that they had no previous knowledge of the figure or its purpose, were told that the problem was to remove three matches from it in such a way that three perfect squares only remained; in other words, to remove three matches in such a way that every match remaining on the table after the three were removed should be a part of a perfect square. No rearranging of the remaining matches was allowed. The subjects were all given exactly the same directions, and were left entirely free to use any method they chose. Removing matches on trial was permitted. Time was counted from the moment the conditions were understood. . . . .

The second ingenuity test was designed to call a pure process of reasoning into play. It consisted of a puzzling mathematical problem, perfectly simple in the computations involved but demanding a somewhat complicated process of reasoning for its solution—a problem in which it was easy to become confused unless all the factors were sharply separated and clearly grasped. The problem was handed written to the subject. He was told that it involved no difficult computations. The process was timed from the moment the problem had been read through. A failure was recorded only in cases in which the subject had worked from forty-five minutes to an hour, and was completely hopeless of getting any solution. The problem was the following: 'A man swimming in a river finds that he can swim three times as fast down stream as up stream. The river flows at the rate of a mile an hour. Find his rate of swimming in still water.' Any solution which could be explained was accepted. A mere stumbling upon the correct answer was not called a solution."

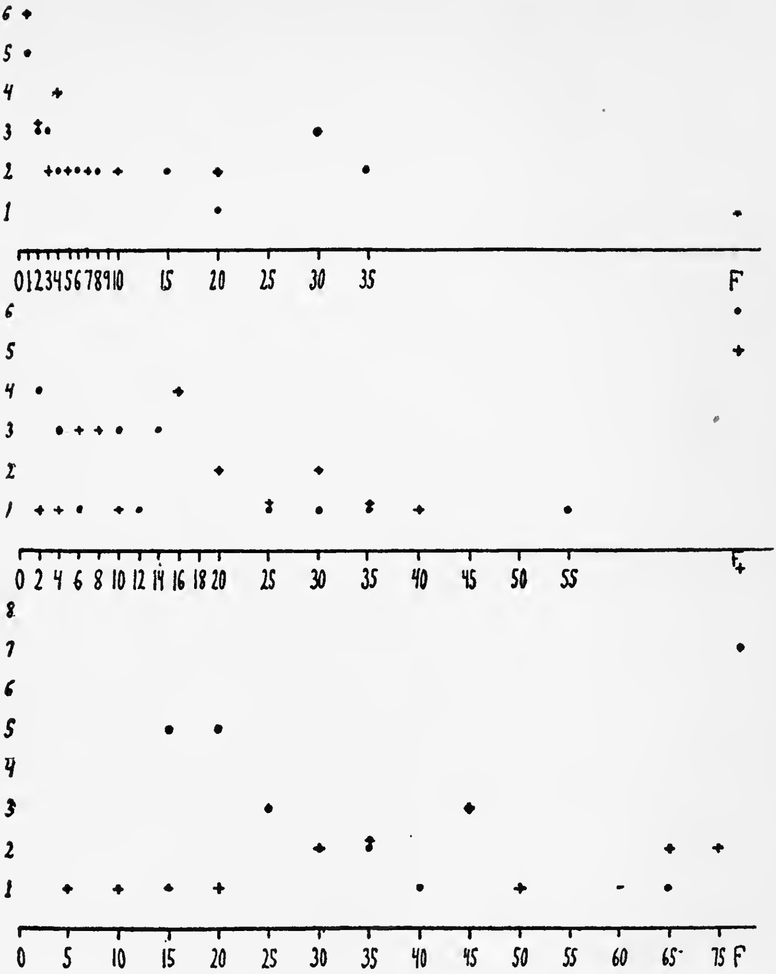
The third problem was to place eight counters on a checkerboard of 64 squares so that no two counters were on the same horizontal, vertical or diagonal row of squares.

The results of these three tests are shown by Dr. Thompson in the curves reproduced in Figures 32, 33 and 34.

Means of Measuring the Differences Between Two Groups

These distribution curves show fully the differences between the men and the women, but they do not show them very conveniently. For convenience in comparing the difference in, say, the second of these tests of ingenuity with the difference in the third or in comparing the difference in any one of them with the difference in memory or accuracy of movement or rate of association, the difference in any one trait should be represented by some one amount.

There are two ways of representing the difference between two groups by one amount. The first is by stating the difference between the two central tendencies. Thus in the second test above the two tables of frequency are as in Table 28. The average is not here a suitable measure of the central tendency



FIGS. 32, 33 and 34. Comparison of men and women in respect to three tests of ingenuity. The horizontal scales are for the time (in minutes) taken to solve the problem. The vertical scales are for the number of individuals. The height of a dot represents the number of men whose times fell within the division of the scale above whose right limit the dot stands. The height of a cross has the same meaning, but for women. Thus (in the middle diagram) 4 men and 1 woman took from 0 to 2 minutes, 3 men and 1 woman from 2 to 4 minutes, 3 women and 1 man from 4 to 6 minutes, in the second test in ingenuity.

FIG. 32 (at the top) records the results in the first test of ingenuity.
FIG. 33 (in the middle) " " " " " second test of ingenuity.
FIG. 34 (at the bottom) " " " " " third test of ingenuity.

since 'failed' cannot be given a numerical value.* There is no clear mode for either men or women. The median is, however, a useful measure here. The median, in the sense of the point on the scale which will have half of the men's records above it and half below, is somewhere between 12 minutes and 14 minutes. Just where it is for these 25 men cannot be stated since we do not know the exact records of the three men who took from 12 to 14 minutes. It would be at the point of the quickest of these three men. The most probable place for him is from 12 to 12⅔ minutes. The midpoint of this place is 12⅓

TABLE 28.

COMPARISON OF MEN AND WOMEN IN INGENUITY—SECOND TEST FREQUENCIES OF DIFFERENT TIMES TAKEN IN SOLVING THE PROBLEM.

	Men		Women
0 to 2 minutes was required by	4	and	1
2 " 4 " " " "	3	"	1
4 " 6 " " " "	1	"	3
6 " 8 " " " "	0	"	3
8 " 10 " " " "	3	"	1
10 " 12 " " " "	1	"	0
12 " 14 " " " "	3	"	0
14 " 16 " " " "	0	"	4
16 " 18 " " " "	0	"	0
18 " 20 " " " "	0	"	2
20 " 25 " " " "	1	"	1
25 " 30 " " " "	1	"	2
30 " 35 " " " "	1	"	1
35 " 40 " " " "	0	"	1
40 " 45 " " " "	0	"	0
45 " 50 " " " "	0	"	0
50 " 55 " " " "	1	"	0
Failed in 45 to 60 minutes	6	"	5

The median man required 12⅓ minutes; the median woman, 15¾ minutes.

minutes. The median for women is the slowest of the 4 records in the 14-16 minute group. The most probable point for this is 15¾ minutes. The median man is thus 3⁵⁄₁₂ minutes quicker than the median woman in this test and requires

* Also for other reasons.

only 8 tenths as long as she. In the third test of ingenuity the median man is the slowest one of the three taking from 20 to 25 minutes, while the median woman is the quicker one of the two taking from 60 to 65 minutes. The most probable points for them are respectively 24.17 and 61.25. The median man is thus 37.08 minutes quicker than the median woman in test 3 and requires only 4 tenths as long.

The second method of representing the difference between two groups by one amount is by stating the per cent of one group that reaches or exceeds a given record made by some one of the other group. Thus in the second test of ingenuity 60 per cent of men reach or exceed (that is, are quicker than) 16 minutes, which is reached or exceeded by 52% of women. 48% of men reach or exceed 12 minutes, which is reached or exceeded by 36% of women. The particular comparison of this sort of most service is *the per cent of group* 1 *reaching or exceeding the median of group* 2. In test 2, 60% of men reach or exceed the median for women. In the case of the third test 69% of men reach or exceed the median for women.

The great advantage gained by comparing groups by the per cent of one group reaching or exceeding the point on the scale that is reached or exceeded by a given per cent of the other group is that results are mutually comparable whatever the traits may be. In place of a list of differences now in time taken, now in amount done, now in quality of this product, now in the amount of that error made, etc., etc., the second method gives a simple list of per cents of men who reach the median for women. Another advantage lies in the fact that this percentile comparison reminds one constantly of the overlapping of the two groups, when such exists.

This second method should, therefore, be used in the statement of sex differences, and may be used exclusively for very small differences and differences measured in ambiguous units such as school and college marks.

Unfortunately, of the investigators who have made mental measurements of men and women, few have realized the need

12

of presenting the distribution of the trait in question for each sex, and still fewer have calculated the per cent of one group passing the point passed by half (or by any other assigned per cent) of the other group. Many of the measurements of sex differences thus remain incommensurate with the rest and are incapable of inclusion in an exact general estimate.

Dr. Thompson does give the entire distributions so that we can summarize the essential features of her results in the table given below (Table 29) :—

TABLE 29.

DIFFERENCES BETWEEN YOUNG MEN AND YOUNG WOMEN IN VARIOUS MENTAL PROCESSES.

In the case of 50 students in the University of Chicago of approximately equal age and academic status the per cent of men reaching or exceeding the median of the women is as follows:

In reaction time	68
Rate of tapping with finger for first 20 seconds	81
Rate of tapping with finger for last 20 seconds of 120*	81
Sorting cards by color; speed	14
Sorting cards by color; accuracy	44
Accuracy in thrust from the shoulder at a target ... approx.	60
Accuracy in free arm drawing of a line within an angle	72
Lowness of threshold for sensations of impact	43
" " " " " " pain	46
" " " " " " taste (the presence of a taste)	34
(recognition of it as sweet, salt, sour or bitter)	34
presence and recognition of sweet and salt	45
of sour and bitter	22
" " " " " " smell (cloves and violet) presence	43
recognition	41
" " " " " " light	62
Range of sensitivity to pitch; upper limit	52
Range of sensitivity to pitch; lower limit	50
" " " " " " lifted weights	66
Delicacy of discrimination of differences in pressure	47

* Six women of the twenty-five could not continue so long as the most easily fatigued man. Two men and two women stopped after 100 seconds. The others tapped for 120 seconds.

Delicacy of discrimination of differences in lifted weights	66
" " " " " " (points on the forearm) crosswise	43
lengthwise	18
" " " " " " area on the skin	61
" " " " " " temperature (at 30 C.)	53
(at 5 C.)	36
(at 45 C.)	52
" " " " " " taste (sweet 64, salt 35, sour 67, bitter 65)	58
" " " " " " smell (cloves 48, violet 51)	50
" " " " " " pitch	44
" " " " " " brightness	78
" " " " " " color (as in tests for color-blindness)	24
" " " " " " area by the eye	56
Memory of nonsense syllables; rate of learning, auditory	32
Memory of nonsense syllables; rate of learning, visual	46
Retentiveness after one week; auditory	51
Retentiveness after one week; visual	43
Quickness in solving ingenuity test 1	46
" " " " " 2	60
" " " " " 3	69
" " " " " 4	68
" " " " " 5	72
General information	50
Information about English literature	31
Information about physics	76

The Results of Measurements of Sex Differences

In examining Table 29 and similar results it will be instructive to have clearly in mind the significance of different per cents of one sex reaching or exceeding the median of the other sex. 50 means of course that the central tendencies of the two sexes are identical.

45 or 55 means the amount of difference shown in Fig. 35.

40 or 60 means the amount of difference shown in Fig. 36.

25 or 75 means the amount of difference shown in Fig. 37.

A per cent of 100 is ambiguous, meaning anything from the amount of difference shown in Fig. 38 to an amount as great as or greater than that shown in Fig. 39. So also a per cent of 0 is ambiguous, meaning possibly a difference as little as that of Fig. 38 (reversed), and possibly a difference as great as or greater than that of Fig. 39 (reversed).

FIG. 35. The amount of difference between two groups when the per cent of one group reaching or exceeding the median of the other group is 45 or 55.

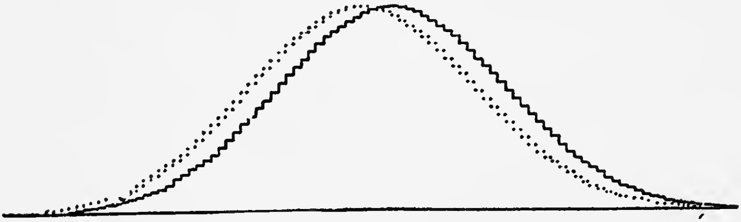


FIG. 36. The amount of difference between two groups when the per cent of one group reaching or exceeding the median of the other group is 40 or 60.

FIG. 37. The amount of difference between two groups when the per cent of one group reaching or exceeding the median of the other group is 25 or 75.

As a matter of fact there is no intellectual ability among those so far measured in which the percentage of males reaching the median for females is as low as 0 or as high as 100. The groups always over-lap to the extent of half the range of one of them, or, more exactly, to the extent of the distance on the scale from the median to one extreme of one group.

Wissler ['01] found, in the case of young men and women

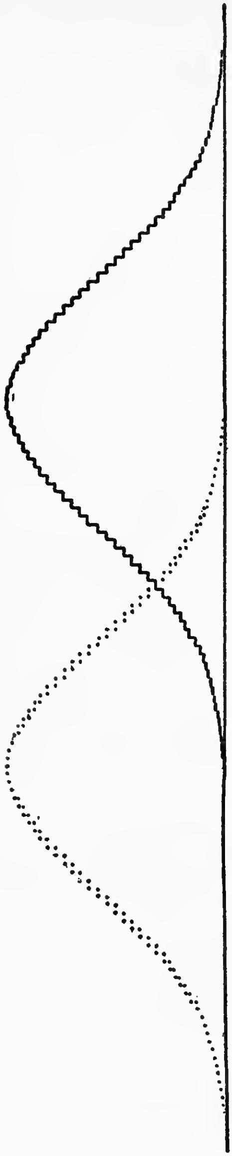


Fig. 38. The least difference possible when the per cent of one group reaching or exceeding the median for the other group is 0 or 100.

Fig. 39.

students in Columbia University, that in fatigue at pressing a spring with thumb and forefinger, in the perception of weight, and in discrimination of points on the skin, there was no appreciable superiority of either sex. Women responded with the judgment of 'painful' to a much less pressure than was required in the case of men, but differences in the standard of 'painful' probably played a large part in the effect. Only 18 per cent of men judged painful as low a pressure as did the median woman. Only 40 per cent of men were as accurate in judging pitch as was the median woman, but in the case of size the figure was 75 per cent. In quickness of reaction time to sound it was 81, but in quickness in marking out A's it was only 32 and in quickness in naming colors it was only 18. In rate of movement about 60 per cent of men equaled or surpassed the median woman, and in rate of association (by a doubtful test) about 70. In memory there was a trifling female superiority.

Gilbert ['94] measured 100 boys and 100 girls of each age from 6 to 17, chosen at random from those in school. He gives simply the medians and the average variations of the individuals therefrom, but it is possible to estimate from his data very closely the percentages of boys reaching or exceeding the median for girls in the case of the traits listed below. The facts are: In the case of boys and girls 8-14 years old (inclusive), the percentage of boys reaching or exceeding the median ability for girls of the same age is:—

In delicacy of sense-discrimination for weight	48
In delicacy of sense-discrimination for color (shades of red)	39
In reaction time	57
In resistance to the size-weight illusion	55
In rate of tapping	64

In the case of boys and girls 15-17 years old (inclusive), the same percentage is:—

In delicacy of sense-discrimination for weight	58
In delicacy of sense-discrimination for color (shades of red)	58
In reaction time	76

In resistance to the size-weight illusion.................. 68
In rate of tapping.................................. 73

I have compared the sexes in the case of various abilities, shown in objective tests and in school marks, with the results shown below. In the case of boys and girls from 8-14 years old (inclusive) the percentage of boys reaching or exceeding the median ability for girls of the same age is:—

In tests of the associative and conceptual processes, such as giving the opposites of words, or the letters that come before given letters in the alphabet, adding, multiplying, and the like.................................. 48
In speed and accuracy in noticing small details, as in marking A's on a sheet of printed capitals, marking words containing certain letters, and the like............... 33
In memory of words for a few seconds (10-30)......... 40
In spelling 33

In the case of boys and girls of the same classes in high schools the same percentage is:—

In tests of the associative and conceptual processes 50 (approx.)
In English (Regents' examination and school mark) 41
In mathematics (Regents' examination and school mark) 57
In Latin (Regents' examination and school mark) 57
In history (Regents' examination and school mark) 60

In the case of college students the same percentage is:—

In English .... 35 (approx.)
In mathematics .... 45 (approx.)
In history and economics.................. 56 (approx.)
In mental science .... 50 (approx.)
In modern languages .... 40 (approx.)

In the case of college students the selection of women is narrower, and probably a little better. The men probably devote less time to their college studies. The students in question were from two state universities in the north-central states.

In an extensive study of the grades received for scholarship in high-school studies (as yet unpublished) Miss Rusk finds the percentage of boys who reach or exceed the station of the median girl to be:—

In English	33
In algebra	41
In geometry	53
In Latin	29
In history	49
In German	34
In chemistry	58
In physics	61

Combining the results of the three investigations of school grades for scholarship in relation to sex, it appears that girls do a bit better in the languages and a bit worse in history, chemistry and physics, there being no discernible difference in the case of mathematics.

The most important characteristic of these differences is their small amount. The individual differences within one sex so enormously outweigh the differences between the sexes in these intellectual and semi-intellectual traits that for practical purposes the sex difference may be disregarded. So far as ability goes, there could hardly be a stupider way to get two groups alike within each group but differing between the groups than to take the two sexes. As is well known, the experiments of the past generation in educating women have shown their equal competence in school work of elementary, secondary and collegiate grade. The present generation's experience is showing the same fact for professional education and business service. The psychologists' measurements lead to the conclusion that this equality of achievement comes from an equality of natural gifts, not from an overstraining of the lesser talents of women.

In detail the measurements show a slight inferiority of the male sex in receptivity or impressibility and a slight superiority in the control of movement and in thought about concrete mechanical situations. Dr. Thompson would attribute the last difference to differences in training, and a charitable male psychologist might so attribute the superior quickness of movement also. The matter is not of great consequence, first because the differences themselves are not, and second because the differences in training, if they exist, are probably due largely to original differences between the interests of the two sexes. If boys by training learn more about the mechanical properties of objects, it is probably because they by nature care more about such learning. It can hardly be maintained seriously that *forced* differences in the training of these 50 students in the University of Chicago or of boys and girls in New Haven and New York—differences in training, that is, apart from the selection of certain training by the children's natures—favored either sex in such a matter as solving an example in arithmetic, or marking out A's, or spelling, or giving the opposites of words.

A vast amount of time could be spent in analyses of the minor differences reported and in argumentation about the reasons for them, for their existence in original nature, and for their relations one with another. It would be largely profitless, however; for no one of these measurements is by itself very reliable and their proper use is only to decide general questions about large differences and about the general extent to which sex is the cause of the mental variations of mankind. They suffice to prove that the sexes are closely alike and that sex can account for only a very small fraction of human mental differences in the abilities listed. They do not suffice to prove the exact nature or amount of the difference in each special trait.

The trivial difference between the central tendency of men and that of women which is the common finding of psychological tests and school experience may seem at variance with the patent fact that, in the great achievements of the world in science, art, invention and management, women have been far excelled by men. One who accepts the equality of typical (i. e., modal) representatives of the two sexes must assume the burden of explaining this great difference in the high ranges of achievement.

The probably true explanation is to be sought in the greater variability within the male sex.* The most gifted men may be superior to the most gifted women even though the average man is equal to or below the average woman, *if men vary widely enough from their central tendency.* Thus in Fig. 40

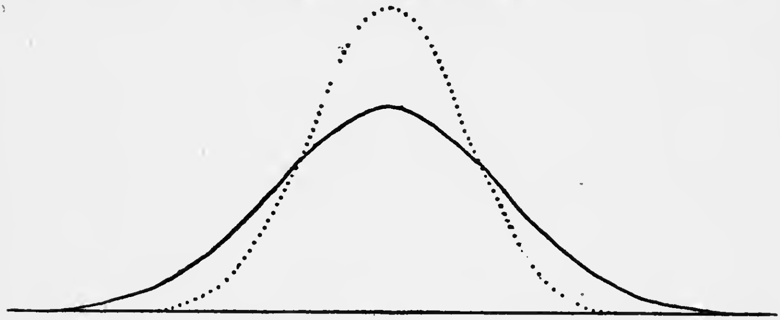


Fig. 40. The continuous line encloses the surface of frequency for men; the dotted line, that for women.

the central tendencies are the same for men and women, but there are two men out of every hundred who are superior to all women. In Fig. 41 only 45 per cent of men reach the median ability for women, but 1 of the 45 is superior to all women.

SEX DIFFERENCES IN VARIABILITY

A difference between the sexes in variability may be of as great significance as a difference in central tendency. This will be clearest if its influence is observed first in one or two imaginary cases. Suppose, for example, that the average position of men on a scale for morality is the same as that for

*It should be obvious that the greater variability of males in the sense of the divergence of individuals from the average or median or mode of their sex implies nothing whatever about the variability of individual men in the sense of the divergence of any man's different 'trials' from his own general average,—in the sense, that is, of the inconstancy of performance of an individual. Men might vary widely *inter se,* but each man might be a very constant performer; women might vary very little from the modal woman, yet each one might vary enormously on different occasions from her average performance or central tendency.

women, and call this amount of morality 20 M. Suppose the average deviation of individual men from 20 M to be 2 M, and the average deviation of women from 20 M to be 3 M. Then the two surfaces of frequency would probably be approxi-

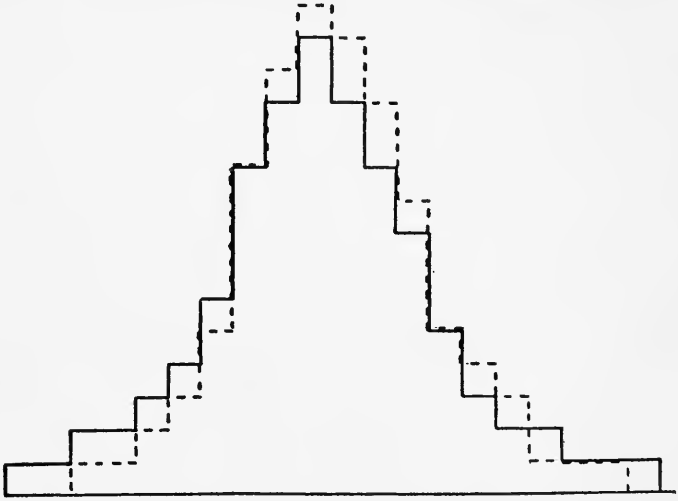


FIG. 41. The continuous line is for men; the broken line, for women

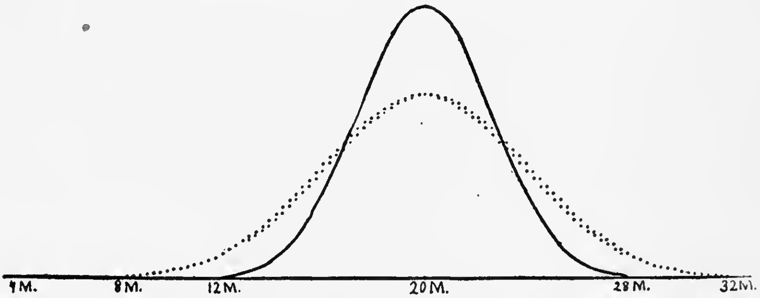


FIG. 42. The status of men and women, the two central tendencies being identical, if the men vary only two-thirds as much as the women.

mately as in Fig. 42. The best man would be about twice as good as the worst man, all men being between about 12 M and 28 M. The limits required to include all the women would, on the contrary, range from about 8 M to 32 M. The best woman would be four times as good as the worst. About two women out of every hundred would be better than the best man; about two would be worse than the worst man.

Thus, though the average morality would be the same, we would have differences of tremendous practical moment. The great acts of honor, philanthropy, nobility and sacrifice would all be due to women. At the same time they would commit all the basest of crimes and iniquities. They would lead in all moral endeavor, but would also fill the jails and dens of wickedness, while the men would present lives of equable, uninteresting mediocrity of both vice and virtue. If the reader will contemplate the practical importance of a similar difference in the variability of the sexes in intelligence, originality, musical talent, piety and other traits, he will see that its measurement is in no wise a matter of merely abstract interest.

In particular, if men differ in intelligence and energy by wider extremes than do women, eminence in and leadership of the world's affairs of whatever sort will inevitably belong oftener to men. They will oftener deserve it. But the greater male variability should result also in a great preponderance of men amongst the most idiotic idiots. Just this seems to be the case. Cattell says ['03, p. 375] in the course of his report on the thousand most noted individuals of the civilized world:—

"I have spoken throughout of eminent men as we lack in English words including both men and women, but as a matter of fact women do not have an important place on the list. They have in all 32 representatives in the thousand. Of these eleven are hereditary sovereigns and eight are eminent through misfortunes, beauty or other circumstances. Belles-lettres and fiction—the only department in which woman has accomplished much—give ten names (of which three are in the first 500) as compared with 72 men. Sappho and Joan d'Arc are the only other women on the list. It is noticeable that with the exception of Sappho—a name associated with certain fine fragments—women have not excelled in poetry or art. Yet these are the departments least dependent on environment and at the same time those in which the environment has been perhaps as favorable for women as for men. Women depart less from the normal than man—a fact that usually holds for the female throughout the animal series; in many closely related species only the males can be readily distinguished. The distribution of women is represented by a narrower bell-shaped curve."

In a study restricted to British genius Ellis ['04, p. 10-11] finds a similar failure of women to reach the extreme of men.

"In the final result my selection yields 975 British men of a high degree of intellectual eminence. The eminent women number 55, being in proportion to the men about 1 to 18.

A slightly lower standard of ability, it would appear, prevails among the women than among the men. On account of the greater rarity of intellectual ability in women, they have often played a large part in the world on the strength of achievements which would not have allowed a man to play a similarly large part. It seemed, again, impossible to exclude various women of powerful and influential personality, though their achievements were not always considerable. I allude to such persons as Hannah More and Mrs. Montague. Even Mrs. Somerville, the only feminine representative of science in my list, could scarcely be included were she not a woman, for she was little more than the accomplished popularizer of scientific results. In one department, and one only, the women seem to be little, if at all, inferior to the men in ability, that is in acting."

It is well known that very marked intellectual weakness is commoner amongst men than amongst women. Two times as many men as women will be found in asylums for idiots and imbeciles; and one and a third times as many will be found by a census including those cases (commonly somewhat less stupid) cared for at home.

In the case of general ability both extremes of both sexes are thus fairly measured for us, but in more specialized traits careful measurements are needed of the comparative variability of men from the typical or 'modal' man and of woman from the 'modal' woman.

Methods of Comparing the Sexes in Respect to Variability

When the two groups are equal in respect to their central tendencies it is easy to compare their variabilities. For instance, in the case of the pressure required to cause a judgment of "painful" the results in Dr. Thompson's investigation ['03] were that the pressure required ranged from 800 to 3600 grams for women and from 800 to 4000 for men. The average deviation of the 25 pressures required for the 25 men from that required for the average or median man was about 960 grams, while that in the case of women was only about 530 grams. Twenty of the twenty-five women are included within a range of 1600 grams, but a range of over 2800 grams is required to include twenty of the twenty-five men. If the men and women were alike in their central tendencies no objection could be raised to comparing their variabilities by the above figures.

But if the men were markedly less sensitive to pain—if the pressure required ranged, for them, from 4000 to 8000, but for women from 800 to 3600—if the median man required 6000 whereas the median woman required only 2200,—then the greater range or greater average deviation of men might not mean a greater real variability. For, an objector could properly say, the average deviation of butterflies from their average in weight is only a small fraction of an ounce whereas the average deviation of men from their average is a hundred or more ounces, yet butterflies really vary more in weight than men do. Only if one man weighs twice as much as another and if one butterfly weighs twice as much as another may the variations be called equal in the two cases, the objector may continue. Equal variability should mean equal *ratios,* not equal *amounts.*

There is obviously much force in this objection, and in the recommendation that a variability around a C. T.* of 20 must be two times as large as a variability around a C. T. of 10 to be properly called equal to it. In certain cases, in fact, this method is demonstrably just. The absolute or gross variation of sermons in length would be much greater than the gross variation in the length of the sentences composing them, but really the latter are much more variable. Again, 22 individ-

*I shall use C.T. as an abbreviation for Central Tendency.

uals worked addition examples, first for forty, then for eighty, then for one hundred and twenty seconds. The scores were, respectively:—

C. T. 9; A. D.* 2.18
C. T. 16; A. D. 3.41
C. T. 24; A. D. 5.18

Now the real variability of these 22 individuals in addition is substantially the same thing in all three tests. They were not two and a half times as much unlike among themselves in the third test as in the first! An Average Deviation of 2.18 around a C. T. of 9 measures the same fact as an Average Deviation of 5.18 around a C. T. of 24. If, instead of the gross variabilities (2.18, 3.41 and 5.18), we use their ratios to the C. T.s (*i. e.,* $\frac{2.18}{9}$, $\frac{3.41}{16}$ and $\frac{5.18}{24}$; or .24, .21 and .22) this sameness in reality is paralleled by the measures. If women tested in addition for 80 seconds showed a C. T. of 16 and an A. D. of 3.4, while men tested for 120 seconds showed a C. T. of 24 and an A. D. of 5.1, it would certainly be absurd to claim therefrom that men were one and a half times as variable as women.

On the other hand, it would be folly to assume that the ratios of the gross variabilities to the corresponding C. T.s are infallibly fair bases for comparing the real variabilities of groups. For instance, tall men vary actually *less* among themselves in stature than do short men. So also men with long middle fingers vary actually less in the length of their middle fingers than do men with short middle fingers.

As a matter of theory the allowance to be made for a difference in central tendency, when groups are compared in variability in a mental trait, should be that obtained by dividing each gross variability by the *square root* of the central tendency rather than that obtained by dividing through by the central tendency itself. In anatomical measurements empirical facts

*A. D. stands for Average, or Mean, Deviation.

support this theoretical expectation. In measurements by scales with arbitrary units, such as school marks, any comparison of groups in respect to variability is treacherous if the groups differ in central tendency. Thus suppose men and women to receive grades in history as follows:—

Grades	40-44	were given	to	0	men	and to		4	women
"	45-49	"	"	0	"	"	"	10	"
"	50-54	"	"	0	"	"	"	23	"
"	55-59	"	"	4	"	"	"	61	"
"	60-64	"	"	10	"	"	"	75	"
"	65-69	"	"	23	"	"	"	61	"
"	70-74	"	"	61	"	"	"	23	"
"	75-79	"	"	75	"	"	"	10	"
"	80-84	"	"	61	"	"	"	4	"
"	85-89	"	"	23	"				
"	90-94	"	"	10	"				
"	95-99	"	"	4	"				

The Average Deviation for men equals that for women. The C. T.s are 62 (60-64) for women and 77 (75-79) for men. The best women get marks about twice as high as the worst women, whereas the best men get marks about one and three-fourths times as high as the worst men. But the variability of men may be really greater than that of women. For the difference called 1 from 85 to 99 may be very much greater than the difference called 1 from 40 to 54. Moreover the fact that the best woman gets marks two times as high as the worst woman tells nothing about how many times as much she knows or how many times as well she can do. The 'times' statement is justifiable only when the 0 of the scale means absolute nothingness of the trait in question. 80 is twice 40, 84 is twice 42, only if it is twice as far from the dividing point between nothing and just barely something on the scale for the trait in question.

In any one mental trait the comparison of men and women in variability may thus be ambiguous, but since the central tendency for males is below about as often as above that for females, the possibility of error tends in the long run almost equally toward exaggeration and toward unfair diminution of male variability. So in a score or more of traits taken at random in respect to this question, any fundamental difference between the sexes in variability should be fairly measured.

The Results of Measurements of Sex Differences in Variability

It is unfortunate that so little information is available for a study of sex differences in the variability of mental traits in the case of individuals over fifteen. Such statistics as I have been able to secure give measures in 26 objective tests, with from 100 to 1,500 individuals, and in 25 records of school marks with from 60 to 1,000 individuals.

The comparisons in the case of reaction time, reaction time with discrimination and choice, and time memory are based on the measurements given by Gilbert ['94]. For the data in spelling, arithmetic and in the r-e and o-n tests I am indebted in part to Messrs. E. L. Earle, W. A. Fox and L. W. Cole.

The nature of the material, which represents measurements taken by different individuals and often with only small groups, makes inference from details unreliable. The data would be slightly more accurate if all records had been reduced to a common month age at least, but this could not be done with the measurements taken by other observers than myself and would involve an amount of labor out of proportion to the increase in accuracy. The main facts that are relevant to our present purpose are as follows: Variability being measured by the per cent which the gross variability is of the central tendency, the sexes differ as shown in Table 30.

If the gross variabilities themselves are used, the ratios for the A test, a-t test, word test, memory of words and spelling are somewhat higher, those for discrimination and reaction time are somewhat lower, while those for the opposite test, time memory, addition, multiplication, and scholarship remain practically the same. The general effect would be to raise the

TABLE 30.

RATIO OF FEMALE TO MALE VARIABILITY.

By Ages.	9	10	11	12	13	14	15	16	17
A test	.86	1.11	1.04	.94	1.08	1.03	1.07		
A-t test	.91	1.05	1.05	1.07	1.35	1.07	.73		
Easy opposites test		.97	.81	1.10	1.24	.89	1.15		
Word test			1.05	.91	.85	.87			
Memory (related words)	.77	1.37	.93	.72					
Memory (unrelated words)	.77	.46	.94	.66	.77	1.28			
Discrimination of length	.75	.80	.81	.98	1.04	.70	.78		
Simple and discriminative reaction time	.98	1.21	.98	.93	1.00	1.11	.83	1.14	1.22
Time memory	.56	.75	1.21	.82	.85	1.27	1.26	.66	1.06
General ratio. Average (9–12; 13–14; 15)			.92		1.025		.97		
Median (9–12; 13–14; 15)			.93		1.035		.95		

The chances are 1 to 1 that the true result will not vary from the one obtained by more than .023 (9-12 yrs.), .04 (13-14 yrs.), .055 (15 yrs.).

By grades.	4	5	6	7	8
R-e and o-n tests	.77	1.19	.97	.82	.85
Spelling	.55	.69	.55	.68	.68
Addition	1.00	.91	1.06	.85	.97
Multiplication				.56	1.15

In a number of tests (six in all) the ratio of first-year high school girls to boys in variability was .975.

In tests in arithmetic (six in all) the ratio of high school girls to boys in variability was .96; in regents' examinations in Latin, English and in history, it was .96; in school marks in eight subjects, it was on the average .86.

In college marks in fourteen different courses the ratios averaged .85.

ratios somewhat since in the particular tests given the central tendency for girls is more often above than below that for boys.

These facts make it extremely probable that, except in the two years nearest the age of puberty for girls,* the male sex is slightly more variable. From the time of puberty for boys to maturity this difference seems to increase rapidly, though the records of marks which support this conclusion are not the best of evidence.

The variability of girls with respect to the age at which any given school grade is reached is less than that of boys. The

*The greater variability of girls in these two years is probably a result of a sex difference in the rate of mental growth.

difference is not necessarily attributable in its entirety to an original difference between the natures of boys and girls. The greater variations of boys toward high ages in particular are probably due in part to the slower maturing of boys, to the greater frequency of temporary withdrawal and to other factors irrespective of an original greater variability. But in so far as boys are found *both younger and older* than girls at entrance to a grade, the evidence of their greater variability in the complex of abilities that determines rate of progress in school is sound. A careful estimate will probably show that in this complex girls are not over ninety-five per cent as variable as boys.

For instance the combined figures for the census of the third-year high school classes (1908) in Detroit, Fall River, Los Angeles, Lowell and Worcester were:—*

Age	13	14	15	16	17	18	19	20 and over.	Total
Boys	3	18	113	237	274	154	69	25	893
Girls	1	12	104	298	265	177	40	10	907

In Chicago (1908) the figures were (the numbers for the 1164 girls being reduced to a basis of 975):—

Age	13	14	15	16	17	18	19	20 and over.	Total
Boys	1	34	165	306	291	127	34	17	975
Girls	1	13	127	371	288	132	32	10	974

In Philadelphia (1908) the figures were (the numbers for the 872 boys being reduced to a basis of 750):—

Age	13	14	15	16	17	18	19	20 and over.	Total
Boys	1	3	42	223	256	168	47	10	750
Girls			31	292	282	111	28	6	750

In New York (1905. Report of Superintendent of Schools, p. 72) the figures were (the numbers for the 1939 girls being reduced to a basis of 1356):—

Age	13	14	15	16	17	18	19	20	21 and over.	Total
Boys	2	37	274	480	382	123	43	12	3	1356
Girls	2	48	300	390	454	134	20	3	5	1356

*For the statistics from which these measurements are computed I am indebted to the U. S. Bureau of Education, through Prof. G. D. Strayer.

On the whole boys are twice as frequent as girls in the youngest and oldest age groups and about one and one-half times as frequent at ages 14 and 19.

Dr. Thompson does not calculate the variability within either sex, nor present the facts on sufficiently fine scales to allow anyone else to do it exactly. I have calculated it as well as may be from the measurements which she gives, with the result that the variability among the 25 women seems on the whole only 93 per cent of that of the men. The difference is closely the same whether the gross variabilities are used directly or are first divided by the corresponding central tendencies or by the square roots of the latter. In reaction time, in the rate of sorting colors, and in memory the 25 women are more variable; in accuracy in hitting a target, in sensory discrimination and in the tests of ingenuity they are almost or quite as variable. But in the majority of tests they are less variable.

Wissler's results ['01] with college students show female variability to be in general about nine-tenths that of males. The number of women measured was, however, only 42, and the ratio of female to male variability differed greatly in the different traits, so that the nine-tenths would, by itself alone, be of no great reliability.

SEX DIFFERENCES IN TRAITS NOT MEASURED OBJECTIVELY

We have now to turn from fairly satisfactory studies of sex differences in sensory, motor and intellectual capacities, to a looser discussion of the life of feeling, action and general achievement. Here objective and precise measurements will seldom be at our service.

There are two studies which do report such differences quantitatively, but the data given are subject, unfortunately, to whatever errors of prejudice or custom teachers, physicians, and German women of intellectual interests make in rating individuals, and to possibly important errors due to the existence in their minds of different standards for the two sexes.

Karl Pearson ['04], in securing data on the resemblances of children of the same parents, had children rated by their teachers for various qualities—as quiet or noisy, shy or self-assertive, and the like. The results in the case where a boy and his sister were both rated are given in Table 31. If taken at their face value they show boys to be somewhat more athletic,

TABLE 31.

RATINGS BY TEACHERS OF BOYS AND THEIR SISTERS IN RESPECT TO VARIOUS TRAITS.

	Boys	Girls
Athletic	291	243
Betwixt	12	9
Non-athletic	131	182
Quiet	440	525
Noisy	313	228
Shy	312	355
Self-assertive	262	218
Self-conscious	380	337
Unself-conscious	277	321
Popular	474	487
Unpopular	81	67
Conscientiousness		
Keen	427	490
Dull	260	197
Temper		
Quick	142	116
Good-natured	490	501
Sullen	72	88
Ability		
Quick-intelligent	131	129
Intelligent	271	302
Slow-intelligent	280	273
Slow	106	110
Slow-dull	49	31
Very dull	23	16
Handwriting		
Very good	51	38
Good	249	313
Moderate	300	274
Poor	111	90
Bad	15	11
Very Bad	3	2

noisier, more self-assertive, more self-conscious, less popular, duller in conscience, quicker-tempered, less sullen, a little duller intellectually, and less efficient in penmanship, in the exact degrees given by Table 31.

They cannot be thus taken unreservedly; for, even in comparing individuals, opinions about the sexes as total groups might be influential and taint the estimates with some measure of current irrational prejudice. This error would probably increase differences beyond their real amounts. Such prejudices, if existing, would work still more insidiously in placing the dividing line between say keenness and dullness of conscience at a different point in the case of boys than was assigned to it in the case of girls. A boy may not have to be really as conscientious or may have to be really more athletic in order to be regarded by current standards as equally conscientious or equally athletic in comparison with his sister. This error would result in reducing differences below their real amounts. However, these measurements are much preferable to general announcements of opinion concerning boys and girls, unless made by specially competent observers.

The greater variability found for males (see Table 31 under Ability and Handwriting) is a sign of the trustworthiness of the data; and the direction of the differences in no case contradicts what little objective evidence exists. So the amounts of difference are worthy of acceptance until a more adequate study is made. They are slight; there is much overlapping of one sex by the other and a far greater range of difference within either sex than between the averages of the two.

On calculating the probable percentages of boys reaching or exceeding the degree of each trait that is reached or exceeded by half of the girls, we have:—

61% of boys are as athletic as or more athletic than the median girl.
62% of boys are as noisy as or more noisy than the median girl.
42% of boys are as shy as or more shy than the median girl.

57% of boys are as self-conscious as or more self-conscious than the median girl.
46% of boys are as popular as or more popular than the median girl.
40% of boys are as conscientious as or more conscientious than the median girl.
56% of boys are as quick-tempered as or more quick-tempered than the median girl.
47% of boys are as intelligent as or more intelligent than the median girl.
43% of boys write as well as or better than the median girl.

Heymans and Wiersma ['06, '07 and '08] studied mental differences of the sexes by means of estimates of individuals made by other individuals who knew them more or less intimately. The report covered 90 topics, some of which included several traits. The individual was graded very coarsely—*e. g.*, as emotional or not emotional; or as a drunkard, an habitual drinker, an occasional drinker, or a total abstainer. Such reports are, as has been noted, inferior evidence, since the person making them may use different standards for men and for women. Thus the same degree of emotionality might be called emotional in the case of a man and not emotional in the case of a woman, or *vice versa*. Moreover when, as often happens, no rating at all is given in a trait, it may be because the condition of the individual to be rated was not known in the case of that trait, or because he was on the dividing line between the two classes (or in the case of a single judgment, like, "Has he mathematical talent?" lacked the specified degree of the quality). Finally, general superstitions about sex differences may affect the ratings of individuals. In the case of the ratings given by women it seems probable that some of the women knew that their ratings were to be used for a study of sex differences. In the case of the ratings given by men this was apparently not so often the case.

On the whole, the results of the ratings, though very inferior to objective measurements, are probably superior to the mere opinions which one could give from reflection on the common facts of life and his own narrow circle of acquaintances. They may at least serve to make the reader critical of whatever such opinions he has.

The authors do not pretend to distinguish, in the case of any of these traits, the differences due to sex itself and the differences due to the difference in the training given to girls. They do, however, give interesting statements of the difference between the present and the previous generation.

Their own conclusions are that the fundamental differences shown by their studies are the greater (1) activity, (2) emotionality, and (3) unselfishness of the female. They consider women to be more impulsive, less efficient intellectually, and more fickle than men as a result of the first two differences mentioned above; to be more gifted in music, acting, conversation and the invention of stories as a result in part of the second difference; and to think well of people and be easily reconciled to them as a result of the third ['07 p. 20].

These conclusions are vague, and the tables of comparison of the sexes which give rise to them are exceedingly long and obscure. The latter take the form of 90 classifications such as:—

	From the reports made almost exclusively by men, the per cents of men and of women were:		From the reports made almost exclusively by women, the per cents were:	
	Men	Women	Men	Women
Emotional............	45	60	49	71
Not emotional........	40	27	40	20

I have therefore calculated from them the probable per cent of men reaching or exceeding the median woman in respect to each trait, counting the ratings by men and those by women as of equal weight. The differences so estimated I have arranged roughly in the order of their magnitude. The largest difference is that:—

Only 15 per cent of men are as much more interested in persons than in things as the median woman is.

The next largest differences are that:—

In accurate and orderly retention of what is read	73%	of men equal or excel the median woman.
In industry	28%	" " " " " " " "
In adroitness in manual work	28%	" " " " " " " "
In love of sedentary games of skill	71%	" " " " " " " "
In emotionality	30%	" " " " " " " "
In temperance in the use of alcoholic drinks	30% (or less)	" " " " " " " "
In independence	70%	" " " " " " " "
In zeal for money making	69%	" " " " " " " "
In desire for change	32%	" " " " " " " "
In impulsiveness	34%	" " " " " " " "
In quickness of recovery from grief	66%	" " " " " " " "

Then come the following:

In activity (of the aimless sort)	36%	of men equal or excel the median woman.
In dissatisfaction with oneself	36%	" " " " " " " "
In religiousness	36%	" " " " " " " "
In excitability	37%	" " " " " " " "
In sympathy	38%	" " " " " " " "
In patience	38%	" " " " " " " "
In love of sports	62%	" " " " " " " "
In humor	61%	" " " " " " " "
In risibility	39%	" " " " " " " "
In talkativeness	40%	" " " " " " " "
In gaiety	40%	" " " " " " " "
In vanity of person	40%	" " " " " " " "

There are very slight differences as follows: men are a little oftener reported as critical, attached to opinions once formed, given to ambitious plans, given to contradiction, sensible, decisive, gifted in mathematics, gifted in literature, specific, of good memories, fond of eating and drinking, fond of distinction, strict, and also easy-going, in discipline with children, kind to subordinates, widely read, and punctual. They are a little less often reported as good-natured, anxious, easily reconciled after anger, insistent on immediate results, good judges of human nature, practically resourceful, narrow, gifted in languages, gifted in music, good observers, thrifty, domineering, kind and careful in discipline with children, active in philanthropic work, demonstrative, honest about money, fond of intercourse with social superiors, timid, well posted about the affairs of acquaintances, polite, attentive, tidy, and courageous in sickness.

In the following traits there is still less difference reported or no difference observable: Trustfulness, tolerance, inconstancy in sympathies, devotion to old memories, quickness in comprehension, superficiality, stupidity, ability in drawing, acting, mimicing, ear for music, patriotism, naturalness, straightforwardness, truthfulness, kindness to animals, snobbishness, courage, and pleasure-seeking.

It is not desirable to comment further on these results of Heymans' and Wiersma's work, until further study by more objective methods has been given to the topic.

It would be desirable in any such study that the sex differences in the instinctive acts, interests, aversions and emotional responses should be studied apart from the differences in similar traits that have been produced by circumstances. Two instincts are worthy of special attention. The most striking difference in instinctive equipment consists in the strength of the fighting instinct in the male and of the nursing instinct in the female. No one will doubt that men are more possessed by the instinct to fight, to be the winner in games and serious contests, than are women; nor that women are more possessed than men by the instinct to nurse, to care for and fuss over others, to relieve, comfort and console. And probably no serious student of human nature will doubt that these are matters of original nature. The out-and-out physical fighting for the sake of combat is pre-eminently a male instinct and the resentment at mastery, the zeal to surpass and the general joy at activity in mental as well as physical matters seem to be closely correlated with it. It has been common to talk of women's "dependence." This is, I am sure, only an awkward name for less resentment at mastery. The actual nursing of the young seems likewise to involve equally unreasoning tendencies to pet, coddle, and "do for" others. The existence of these two instincts has been long recognized by literature and common knowledge, but their importance in causing differences in the general activities of the sexes has not. The fighting instinct is in fact the cause of a very large amount of the world's intellectual endeavor. The financier does not think merely for money nor the scientist for truth nor the theologian to save souls. Their intellectual efforts are aimed in great measure to outdo the other man, to subdue nature, to conquer assent. The maternal instinct in its turn is the chief source of woman's superiorities in the moral life. The virtues in which she excels are not so much due to either any general moral superiority or any set of special moral talents as to her original impulses to relieve, comfort and console.

Training undoubtedly accentuates these inborn differences since boys play more with boys and are trained more by men, the opposite holding with girls. A reversal of training by which girls would be surrounded by the social milieu now affecting boys would, as we often see in isolated cases, lessen the sex difference. But we may be sure that if we should keep the environment of boys and girls absolutely similar these instincts would produce sure and important differences between the mental and moral activities of boys and girls.

Since these differences in instinctive equipment are true causes it seems wise not to invoke other less probable traits to account for any fact which these seem fairly adequate to explain. For instance, if the intellectual achievement of men was found to be superior to that of women we could explain it either by the indirect effect of physical strength and bodily fitness or by an actual difference in intellect or by the zeal and activity due to the fighting instinct. Our rule would be to exhaust first the influence of the known physical differences and second the influence of the instinct in question. Only if these were inadequate should we resort to the hypothetical cause of differences in purely intellectual caliber.

Havelock Ellis ['94, '04] chooses as general sex differences the less variability and the greater affectability and primitiveness of the female mind. The first point has been discussed fully. By affectability he means not only greater impressibility by and responsiveness to stimuli of all sorts, but also less inhibition of the emotions and other instinctive reactions. The fact seems indubitable though its exact amount can not be even roughly estimated. Not only the superiority in tests of perceptual power and the greater suggestibility which we have noted, but also the relative frequency of dreams, trance states and emotional outbreaks and the common differences between our treatment of the men and of the women with whom we are associated, witness to it. In his evidence for the discussions of the primitive nature of women Mr. Ellis seems to have physique in view primarily. How far women resemble uncivilized races and children in mental make-up is, to me at least, not at all clear.

The same author emphasizes, as so many others have done, the fact of female dependence or lack of aggressiveness in intellect. The qualities that we call original, constructive, organizing and critical are ill defined and comparisons are hard to arrange because men and women have devoted the active power of the intellect to such different fields. Comparison of the most eminent representatives from both sexes is obviously unfair in so far as men are more variable.

If we are to believe the novelists and playwrights, women are more concerned with their own feelings and personalities than men, are emotionally more subjective. This is not inconsistent with the existence of greater sympathy of the motherly sort, nor with the possibly superior gifts of men in the examination and intellectual manipulation of subjective conditions. An interesting bit of evidence supports the conventional view of fiction. Many people carry on as a systematic day dream a continued story in which they figure and which possesses its interest from the chance it gives to think pleasantly of oneself. According to Learoyd ['96] three and a half times as many women as men do this (46.7 per cent and 23.5 per cent).

On the whole the differences reported in the case of the less easily measurable features of intellect, character and behavior are of the same order of magnitude as those found in objective tests. They do not require any amendment of the general rule that sex is the cause of only a small fraction of the differences between individuals. The differences of men from men and of women from women are nearly as great as the differences between men and women.

CHAPTER X

The Influence of Remote Ancestry or Race

Men are mentally like one another and unlike dogs or horses because men spring from a presumably common remote ancestry which was not the ancestry of dogs and horses. Men, dogs and horses are more alike mentally than men, dogs, horses, earthworms and clams are, because presumably men, dogs and horses spring from a common ancestry which was not the ancestry of earthworms or clams. Certain men, for example the American Indians, springing from a common ancestry which was not the ancestry of Europeans, may be expected to be mentally more like one another than like Europeans, if their common ancestry differed mentally from that of Europeans.

A distinct race is a group of men who to a considerable extent have in common the same remote ancestry, its present descendants being to a considerable extent confined to that group. The more they all hark back to just the same ancestry, and the more exclusively they represent the present product of this ancestry, the more distinct a race they will be.

A race that is thus distinct in ancestry is commonly distinct in some physical traits also. If its remote ancestry differed from the remote ancestry of other races in mental traits, as it probably did to at least some slight extent, the race has a probability of differing from other races in these traits. So also if the race has differed from other races in the nature or amount of selection, natural or artificial, on the basis of mental traits. An individual may thus, by original nature, possess certain racial mental tendencies. His position on the scale for any mental trait may be due in part to his membership in a certain race,—that is, to his origin from a certain remote ancestry.

The influence of remote ancestry cannot, however, be isolated for measurement at all perfectly. The pedigrees of human stocks, at least of the modern civilized stocks, are not clear enough, and the influence of similarity in remote ancestry is hard to distinguish from that of similarity in training. Many of the mental similarities of an Indian to Indians and of his differences from Anglo-Saxons disappear if he happens to be adopted and brought up as an Anglo-Saxon.

So, though the best way to think of the problem is to picture the hereditary relations of all men and to measure the original likenesses within the same strains and the differences between strains, apart from all influences of training, the facts at hand do not so group themselves. The facts are measurements of the differences between groups which are distinct to an unknown degree in traits which are influenced by training to an unknown degree. Any conclusion will depend upon one's estimates of these two unknown quantities.

I shall report one sample study of the topic in the case, of great importance to American education, of differences between whites and negroes in scholarship in the high school. I shall then summarize the results of other measurements of racial mental differences; and, lastly, describe the attitude which science recommends toward the apparently original differences in intellect, character and temperament which have not been measured.

A SAMPLE STUDY OF RACIAL DIFFERENCES

Mayo ['13] secured the academic records of 150 negroes* who entered the high schools of New York City since 1902.

*A negro being defined as an individual reported as a negro by school officers. Mulattoes are of course frequent.

For each such record he got a white pupil's* record selected under the same conditions. It is impossible to tell exactly whether and how far the two groups of pupils thus taken represent dissimilar samplings of the two total groups, negroes and whites in New York City. In my opinion the samplings are closely similar. There are no measurements of the extent to which residence in New York selects the more scholarly of negroes from the country, or of the extent to which entrance to high school selects differently from the negroes in New York than from the whites. In general, selection by entrance to the public high schools is narrow but democratic; and in Mr. Mayo's opinion and my own the high school gets a somewhat, but not much, higher selection from the colored than from the white youth. That is, in our opinion, the superiority of the colored in high school to the colored outside is greater, but not much greater, than the superiority of the whites in high school to the whites outside.

Whatever be the difference in the selection of the two groups, colored beginners in high schools in New York City differ from whites in their careers there as follows:

(1) On the average they are seven months older, only 36 per cent of them being as young as the median white.

(2) They continue in the high school longer.

(3) In achievement in the different studies they are somewhat, but not very much, inferior. The general tendency is for only three-tenths of them to reach the median record for whites.

(4) The difference is greatest in the case of English, in which only 24 per cent of the colored pupils reach or exceed the median for whites.

Table 32 gives a sample of Mr. Mayo's measurements of the differences. Fig. 43 presents the facts graphically.

* A white pupil being defined as an individual reported as such by school officers. There may in rare cases have been some slight mixture of negro blood.

TABLE 32.

White and Colored Pupils Compared in Scholarship in New York City High Schools.

Median of all marks in the first trials (that is, excluding marks for courses repeated after failure)	Number of white pupils	Number of colored pupils
24—25	..	1
26—27	..	..
28—29	..	..
30—31	..	..
32—33	1	..
34—35	1	..
36—37	..	..
38—39	..	..
40—41	1	5
42—43	2	..
44—45	..	2
46—47	..	2
48—49	..	1
50—51	4	5
52—53	4	6
54—55	1	9
56—57	2	7
58—59	6	12
60—61	18	17
62—63	21	22
64—65	13	16
66—67	12	13
68—69	11	15
70—71	10	2
72—73	16	1
74—75	8	3
76—77	3	3
78—79	4	1
80—81	5	..
82—83	2	3
84—85	2	1
86—87	..	..
88—89	..	1
90—91	2	..
92—93	1	1
Median	**66**	**62**

Mr. Mayo also measured the differences in variability, as far as he could from the arbitrary measures afforded by school marks. For the details the reader must turn to his report, one sample table and the general drift of the results being all that can be presented here. The records of colored pupils were perhaps a very little less variable than the whites. Thus, in the score for total scholarship (Table 32), 80 per cent of the colored pupils are included within a range of 19½ points on the scholarship scale, while to include 80 per cent of the white

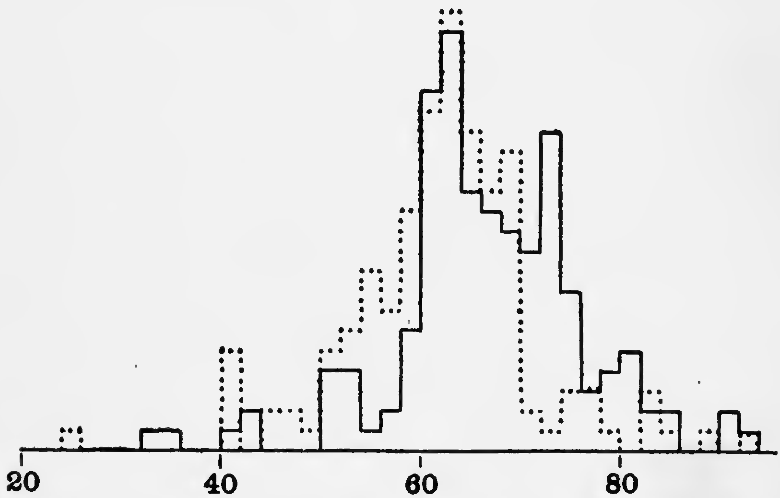


FIG. 43. Comparison of White Pupils (continuous line) and Colored Pupils (dotted line) in respect to Scholarship in the High School. The horizontal scale is for the median of all marks obtained by an individual except those obtained in courses repeated because of failure. The marks in these schools are on a 0 — 100 scale.

pupils requires a range of 20 points. The corresponding figures for the inclusion of 60 per cent are 14½ and 14 points. The figures for the inclusion of 50 per cent are 9 and 10½. The figures for the inclusion of 90 per cent (or, to be exact, 90.7 per cent) are 33 and 31. The comparison in variability is, as in the case of the sexes, of great practical importance. The ability of a hundred of its most gifted representatives often counts more for a nation's or a race's welfare than the ability of a million of its mediocrities. The fact that the colored pupils, though clustered more closely at the center, range to nearly or quite as high grades as do the white is thus noteworthy.

THE RESULTS OF MEASUREMENTS OF RACIAL MENTAL DIFFERENCES

In summarizing the measurements of racial mental differences I shall in the main quote the admirable account by Woodworth ['10].

The different races of men have been compared by objective measurements in respect to the keenness of the senses, but in few traits besides. The reports of travelers gave rise to the doctrine that primitive races excelled modern Europeans in powers of vision, hearing and smell. Skepticism concerning this doctrine has led to many measurements by anthropologists.

"Ranke on testing natives of Brazil, a race notable for its feats of vision, found that their ability to discern the position of a letter or similar character at a distance, though good, was not remarkable, but fell within the range of European powers. The steppe-dwelling Kalmuks, also renowned for distant vision, being able to detect the dust of a herd of cattle at a greater distance with the naked eye than a European could with a telescope, have also been examined; and their acuity was indeed found to be very high, averaging considerably above that of Europeans; yet only one or two out of the forty individuals tested exceeded the European record, while the great majority fell within the range of good European eyes. Much the same result has been obtained from Arabs, Egyptians and quite a variety of peoples. Among the most reliable results are those of Rivers on a wholly unselected Papuan population. He found no very exceptional individual among 115 tested, yet the average was somewhat better than that of Europeans. I had myself, through the kindness of Dr. McGee, the opportunity of testing individuals from quite a variety of races at the St. Louis Fair in 1904, and my results agree closely with those already cited, though I did not find any cases of very exceptional powers among about 300 individuals. There were a number who exceeded the best of the 200 whites whom I also tested under the same conditions, but none who exceeded or equaled the record of a few individuals who have been found in the German army. Indians and Filipinos ranked highest, averaging about 10 per cent better than whites, when all individuals of really defective vision were excluded. The amount of overlapping is indicated by stating that 65-75 per cent of Indians and Filipinos exceeded the average for whites." ['10, p. 175]

There are racial differences in hearing, as tested by the ticking of a watch or a click artificially made. The Papuans were found to be inferior to Europeans. Bruner ['08, p. 92] found that only five per cent of Filipinos equalled or exceeded the median white American. The per cent for Indians with more or less school training was 38. Of the 18 Patagonians, Ainus and reputed Pygmies, none equalled the median white American. These differences are probably due in large measure and possibly *in toto* to the greater cleanliness, freedom from injury to the ear, and special training in hearing a click transmitted by a telephone.

Of the sense of touch and the sense of pain Woodworth says:

"The sense of touch has been little examined. McDougall found among the Papuans a number with extremely fine powers of discrimination by the skin. The difference between two points and one could be told by these individuals even when the two points were brought very close together; on the average, the Papuans tested excelled Europeans considerably in this test. On the other hand, Indians and Filipinos, and a few Africans and Ainu, tested in the same manner, seem not to differ perceptibly from whites.

The pain sense is a matter of some interest, because of the fortitude or stolidity displayed by some races towards physical suffering. It may be, and has been conjectured, that the sense for pain is blunt in these races, as it is known to be in some individuals who have allowed themselves to be burned without flinching, and performed other feats of fortitude. The pain sense is tested by applying gradually increasing pressure to some portion of the skin, and requiring the person tested to indicate when he first begins to feel pain. Now as a matter of fact, the results of McDougall on the Papuans, and those of

Dr. Bruner and myself on Indians, Filipinos, Africans and Ainu, are in close agreement on this point. Greater pressure on the skin is needed to produce pain in each of these races than in whites. This is the average result, but in this test the distribution of the cases is specially important. Though most whites feel pain at or about a certain small pressure, there is quite a respectable minority who give no sign till much higher pressures are reached, their results corresponding very closely to those of the majority of Indians. And similarly, a minority of Indians feel pain at much lower pressures than the bulk of their fellows, falling into the ranks of the white man. In each group, the distribution is bimodal, or aggregated about two points instead of one; but whites are principally aggregated about the lower center, and Indians and other races about the higher center. Introspection comes to our aid in explaining this anomaly, for it shows that there is some difficulty in telling just when the pressure becomes painful. If one is satisfied with slight discomfort, a moderate pressure will be enough; but if a sharp twinge is demanded, the pressure must be considerably increased. Most whites, under the conditions of the test, are satisfied with slight discomfort, while my impression in watching the Indians was that they were waiting to be really hurt. The racial difference would accordingly be one in the conception of pain, or in understanding the test, rather than in the pain sense.

On the whole, the keenness of the senses seems to be about on a par in the various races of mankind." ['10, pp. 176-177]

With respect to racial differences in speed of brain action and in certain objective, though somewhat trivial, tests of intellect, Woodworth says:

"Some interest attaches to tests of the speed of simple mental and motor performances, since though the mental process is very simple, some indication may be afforded of the speed of brain action. The reaction time test has been measured on representatives of a few races, with the general result that the time consumed is about the same in widely different groups. The familiar 'tapping test,' which measures the rate at which the brain can at will discharge a series of impulses to the same muscle, was tried at St. Louis on a wide variety of folk, without disclosing marked differences between groups. The differences were somewhat greater when the movement, besides being rapid, had to be accurate in aim. The Eskimos excelled all others in this latter test, while the poorest record was made by the Patagonians and the Cocopa Indians—which groups were, however, represented by only a few individuals. The Filipinos, who were very fully represented, seemed undeniably superior to whites in this test, though, of course, with plenty of overlapping. . . . . .

"Equitable tests of the distinctly intellectual processes are hard to devise, since much depends on the familiarity of the material used. Few tests of this nature have as yet been attempted on different races.

"There are a number of illusions and constant errors of judgment which are well known in the psychological laboratory, and which seem to depend, not on peculiarities of the sense organs, but on quirks and twists in the process of judgment. A few of these have been made the matter of comparative tests, with the result that peoples of widely different cultures are subject to the same errors, and in about the same degree. There is an illusion which occurs when an object, which looks heavier than it is, is lifted by the hand; it then feels, not only lighter than it looks, but even lighter than it really is. The contrast between the look and the feel of the thing plays havoc with the judgment. Women are, on the average, more subject to this illusion than men. The amount of this illusion has been measured in several peoples, and found to be, with one or two exceptions, about the same in all. Certain visual illusions, in which the apparent length or direction of a line is greatly altered by the neighborhood of other lines, have similarly been found present in all races tested, and to about the same degree. As far as they go, these results tend to show that simple sorts of judgment, being subject to the same disturbances, proceed in the same manner among various peoples; so that the similarity of the races in mental processes extends at least one step beyond sensation.

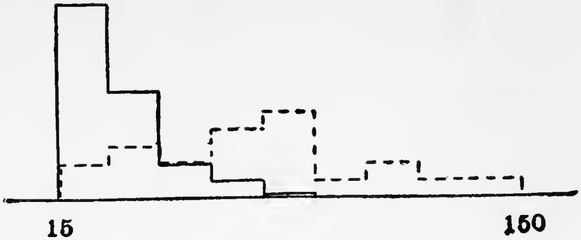
"The mere fact that members of the inferior races are suitable subjects for psychological tests and experiments is of some value in appraising their mentality. Rivers and his collaborators approached the natives of Torres Straits with some misgivings, fearing that they would not possess the necessary powers of sustained concentration. Elaborate introspections, indeed, they did not secure from these people, but, in any experiment that called for straightforward observation, they found them admirable subjects for the psychologist. Locating the blind spot, and other observations with indirect vision, which are usually accounted a strain on the attention, were successfully performed. If tests are put in such form as to appeal to the interests of the primitive man, he can be relied on for sustained attention. Statements sometimes met with to the effect that such and such a tribe is deficient in power of attention, because, when the visitor began to quiz them on matters of linguistics, etc., they complained of headache and ran away, sound a bit naive. Much the same observations could be reported by college professors, regarding the natives gathered in their class rooms.

"A good test for intelligence would be much appreciated by the comparative psychologist, since, in spite of equal standing in such rudimentary matters as the senses and bodily movement, attention and the simpler sorts of judgment, it might still be that great differences in mental efficiency existed between different groups of men. Probably no single test could do justice to so complex a trait as intelligence. Two important features of intelligent action are quickness in seizing the key to a novel situation, and firmness in limiting activity to the right direction, and suppressing acts which are obviously useless for the purpose in hand. A simple test which calls for these qualities is the so-called 'form test.' There are a number of blocks of different shapes, and a board with holes to match the blocks. The blocks and board are placed before a person, and he is told to put the blocks in the holes in the shortest possible time. The key to the situation is here the matching of blocks and holes by their shape; and the part of intelligence is to hold firmly to this obvious necessity, wasting no time in trying to force a round block into a square hole. The demand on intelligence certainly seems slight enough; and the test would probably not differentiate between a Newton and you or me; but it does suffice to catch the feeble-minded, the young child, or the chimpanzee, as any of these is likely to fail altogether, or at least to waste much time in random moves and vain efforts. This test was tried on representatives of several races and considerable differences appeared. As between whites, Indians, Eskimos, Ainus, Filipinos and Singhalese, the average differences were small, and much overlapping occurred. As between these groups, however, and the Igorot and Negrito from the Philippines and a few reputed Pygmies from the

Congo, the average differences were great, and the overlapping was small. Another rather similar test for intelligence, which was tried on some of these groups, gave them the same relative rank. The results of the test agreed closely with the general impression left on the minds of the experimenters by considerable association with the people tested. And, finally, the relative size of the cranium, as indicated, roughly, by the product of its three external dimensions, agreed closely in these groups with their appearance of intelligence, and with their standing in the form test. If the results could be taken at their face value, they would indicate differences of intelligence between races, giving such groups as the Pygmy and Negrito a low station as compared with most of mankind. The fairness of the test is not, however, beyond question; it may have been of a more unfamiliar sort to these wild hunting folk than to more settled groups. This crumb is, at any rate, about all the testing psychologist has yet to offer on the question of racial differences in intelligence." ['10, pp. 179-181]

The difference between civilized whites and Negritos mentioned by Professor Woodworth is shown graphically in Fig. 44 for the first trial in placing the blocks and in Fig. 45 for the third trial. In the first trial only 9½ per cent of the Negritos were as quick as the median white; in the third trial no one of them was, the best individual of the twenty-two just not reaching the speed of the median white. The Negritos also made many more errors. The reputed Pygmies were still less capable than the Negritos. The Pygmies apparently did not do as well in this test as the so-called 'feeble-minded' and 'higher grade imbeciles' confined in state asylums in Massachusetts and Connecticut.

Havelock Ellis ['04] studied the relation of race to amount and kind of achievement in the case of the 1030 most eminent British careers from the 4th to the 19th century. He verifies the opinion that the Scotch have a larger percentage of men of great intellect than the English, Welsh or Irish. The most marked difference which he finds is between the Scotch and Irish in the case of ability in Science and in Acting. Of the 120 men of science, 21 were Scotch and only 1 Irish, while of



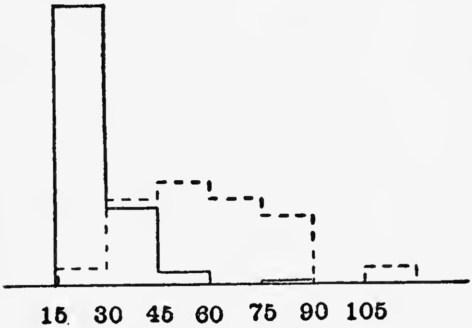


Fig. 44 (upper diagram). Comparison of whites (continuous lines) and Negritos (dotted lines) in respect to time taken to put variously shaped blocks in holes to match. The horizontal scale is for time in seconds, first trial.

Fig. 45 (lower diagram). As in Fig. 44, except that the records in the third trial were used.

the 42 actors 6 were Irish and none were Scotch. He also finds signs of differences in the character of the scientific work done by men from different sections representing somewhat different racial stocks. He writes:

"Psychologically it is not difficult to detect a distinct character in English scientific genius, according as it springs from the Anglo-Danish district or the East-Anglian focus or the southwestern focus, although I am not aware that this has been pointed out before. The Anglo-Danish district may here be fairly put first, not only on account of the large number of scientific men it has wholly or in part produced, but also on account of the very high eminence of some among them. The Anglo-Dane appears to possess an aptitude for mathematics which is not shared by the native of any other English district as a whole, and it is in the exact sciences that the Anglo-Dane triumphs.* Newton is the supreme figure of Anglo-Danish science; it will be noted that he belongs to the East-Anglian border, and by his mother is claimed by Rutland, a little county which, I am inclined to think, really belongs psychologically and perhaps ethnologically to East Anglia. The combination of the Anglo-Dane and the East Anglian seems highly favorable to scientific aptitude; the abstracting tendency of the Anglo-Dane, and the exaggerated independence of his character, with the difficulty he finds in taking any other point of view than his own, are happily tempered by the more cautious and flexible mind of the East Anglian. Darwin (who also belonged to the Welsh Border) belonged in part, like Newton, to the East Anglian border of the Anglo-Danish district, and also (somewhat remotely) to Norfolk, a county which contains many Danish elements. The science of the Anglo-Danish district is not exclusively mathematical, and geology especially owes much to the Anglo-Dane; it will be remembered that geology was one of the first sciences to attract Darwin.

"The East Anglian is in scientific matters drawn to the concrete, and shows little or no mathematical aptitude. He is a natural historian in the widest sense. He delights in the patient collection of facts, and seeks to sift, describe, coordinate, and classify them. In his hands science becomes almost an art. Gilbert illustrates East Anglian scientific methods in the inorganic world, Ray in the organic, and Francis Bacon, though he cannot himself be classed among men of science, has in the Novum Organum and elsewhere presented a picture of scientific method as it most naturally appears to the East Anglian mind.

"It is not easy to see anything specific or definitely Brythonic in the scientific activities of the Welsh Border. At most it may be said that there is some tendency for science here to take on a technological character and to become associated with the artistic crafts. The scientific men found here often belong only in part to the district, and many of them seem to possess the psychological character of the southwestern focus.

"The scientific characters of the southwestern focus are quite clear, and definitely distinct from those of either the Anglo-Danish district or the East Anglian focus. What we

*The mathematical tendencies of Cambridge are due to the fact that Cambridge drains the ability of nearly the whole Anglo-Danish district.

find here is the mechanical impulse, and more especially the physiological temper, the instinct to seek out the driving forces of vital phenomena. It is on this account that Harvey, though of Kentish family, may be said to belong psychologically to this focus, as also Stephen Hales, though he belonged partly to Kent and partly to East Anglia. The great scientific physicians belong here (the surgeons are largely East Anglian), with Sydenham at the head and Glisson. Huxley, again, is a typical figure. Inventors are numerous, for the scientific men of this region have frequently been enamoured of practical problems, and just as they have been pioneers in the physical world, so in science they have sought rather to make discoveries than to formulate laws. Thus in astronomy we have Adams, and one of the greatest and most typical scientific men of this region was Thomas Young." ['04, pp. 68-71]

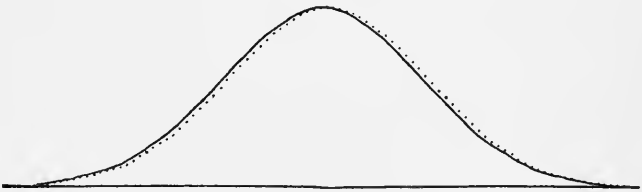


Fig. 46.

These last differences are not measured, but if they exist at all they are surely very slight; for even the most striking difference, that between the Scotch and Irish in science and in acting, is really a very small difference. Even supposing circumstances of religion, education and the like to have had no part in it and so attributing the whole of the difference found to race, the facts found could be explained by supposing the two races to differ in the capacity for science by much less than the amount shown in Fig. 46. Similarly in the case of acting. A very slight difference between two groups in central tendency may (and will, except for contrary influences from differences in variability) make an enormous difference between the percentages from the two groups that possess a very high degree of the trait in question. Indeed I have quoted Ellis's conjectures chiefly as a sample of how, on the one hand, a striking difference between extreme representatives may mislead one and of how, on the other hand, a very small general difference between two races may by its effect in producing many more men of very high ability, advance the social condition of the favored race.

THE INTERPRETATION OF THE DIFFERENCES BETWEEN ONE RACE AND ANOTHER IN ACHIEVEMENT

The moderation of these findings by psychologists is in striking contrast with the first and common impression made by the history and present status of different races of men. The modern European, who can kill a hundred of a race in an hour, or buy up their entire property with the results from one day of his labor, or get any of the results one of them desires in a tenth of the time that they take, seems to be far more capable than they. The Chinese, who resists all that we think he ought to crave, seems obviously to have a temperament radically different from ours. Large, glaring differences mark the achievements of races and seem to need large differences in nature as their origins. The most noticeable fact about the races of men seems to be their great mental variety. Under the deliberate scrutiny of actual measurements, however, what seemed to be large differences shrink to five or ten per cent, and what seemed to be wide gaps are bridged. No two races have been measured which do not overlap mentally whatever be the trait measured.

The first thing to note in respect to the apparent conflict between common observation and precise experiments is that the two have not measured the same traits. Common observation of the African and the European, for example, decides that the latter is superior in intellect, enterprise and self-reliance. Even when experiments show him to be approximately equal in sense keenness, resistance to the size-weight illusion, or putting blocks in holes that fit them, the claims of common observation are not necessarily denied. The nature and amount of race differences in such traits as intellect, enterprise and self-reliance cannot be inferred from the amount of difference found in sensory or sensori-motor traits, but must be studied directly. We do not know just what the symptoms of intellect are, but apparently quickness and accuracy in making purely mental associations, ability to respond to parts or elements of situations which cannot be abstracted in reality but only in thought, the consequent power to devise new responses to old situations, and a marked development of the instinctive satisfaction in thought for its own sake are leading ones. Measurements of these traits in different races are much needed. In the second place, two races need not be equally gifted because each is equally well adapted to its environment, if the second race has by superior enterprise sought out or created a more exacting but also more remunerative environment. The Bushman may count all that he needs to count, but to put oneself in a position that needs algebra and the calculus may itself be a symptom of superiority. So the complex of qualities which is called enterprise remains largely untouched by the psychologist's tests. The very fact that a certain test seems to be unfair to the Bushman may be evidence of his inferiority.

A third fact for consideration is that although the most rigorous thinkers amongst anthropologists are skeptical concerning original mental racial differences, the general body of scientific opinion is by no means fully agreed with them. Francis Galton ['69, '92] in a well known chapter on "The Comparative Worth of Different Races" declares that, taking negroes on their own intellectual ground, they still are inferior to Europeans by about one-eighth of the difference between say Aristotle and the lowest idiot. That is, he considers the two groups to differ approximately as shown in Fig. 47. His argument is:

"Thirdly, we may compare, but with much caution, the relative position of negroes in their native country with that of the travellers who visit them. The latter, no doubt, bring with them the knowledge current in civilized lands, but that is an advantage of less importance than we are apt to suppose. A native chief has as good an education in the art of ruling men as can be desired; he is continually exercised in personal government, and usually maintains his place by the ascendency of his character, shown every day over his subjects and rivals. A traveller in wild countries also fills, to a certain degree, the position of a commander, and has to confront native chiefs at every inhabited place. The result is familiar enough—the white traveller almost invariably holds his own in their presence. It is seldom that we hear of a white traveller meeting with a black chief whom he feels to be the better man. I have often discussed this subject with competent persons, and can only recall a few cases of the inferiority of the white man,—certainly not more than might be ascribed to an average

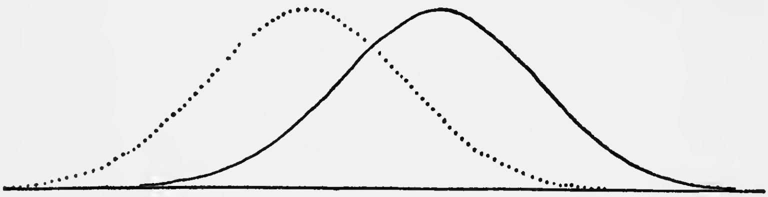


FIG. 47. The Original Difference between Europeans (continuous line) and Negroes (dotted line) in Intellectual Ability, according to Galton.

actual difference of three grades, of which one may be due to relative demerits of native education, and the remaining two to a difference in natural gifts.

"Fourthly, the number among the negroes of those whom we should call half-witted men is very large. Every book alluding to negro servants in America is full of instances. I was myself much impressed by this fact during my travels in Africa. The mistakes the negroes made in their own matters were so childish, stupid, and simpleton-like, as frequently to make me ashamed of my own species. I do not think it any exaggeration to say, that their c is as low as our e, which would be a difference of two grades, as before. I have no information as to actual idiocy among the negroes—I mean, of course, of that class of idiocy which is not due to disease." [Hereditary Genius, 2nd Edition, '92, p. 327 f.]

On the other hand, common observation does not as a rule observe mental traits, but only certain indirect consequences of them which it is likely to misinterpret. It observes customs, not moral capacity; habits, not energy; knowledge, not intellect. But, obviously, the habits and knowledge possessed by a race do not measure its present original nature. Its habits and knowledge, its "civilization" or "culture" are in the main due to the original nature of men long dead and have come to it by training. The origination of advances in civilization is a measure of ability, but the abilities that have originated them have probably been confined to a very few men. A race that originated none of them may now possess them all. Even if a race has been completely isolated, its civilization has been originated by only a few of its members; and the chance of men of great gifts being born is the result not only of the central tendency of a race and its variability, but also of its size. Other things being equal, there is a far greater chance of the birth of a man of great ability in a tribe of a million than in one of a thousand. Since one such man may add to the knowledge and improve the habits of the entire group regardless of its size, civilization will progress more rapidly in large than in small groups, in a condition of isolation.

The civilized races have not remained isolated and have got most of their civilization from without. Of ten equally gifted races in perfect intercourse each will originate only one-tenth of what it gets. The original nature of the Germans of to-day is not much different from that of their ancestors in the time of Tacitus, and their progress in the meantime is not properly theirs, but that of the European world and its American colony, each of whose racial stocks has added something to a common fund.

Again, the civilization—the habits and customs—of a race need not be in a direct proportion to its intellect, even if entirely caused by it. A very slight difference in intellect might give one race supremacy over another, enable it to condemn the other to servitude and so free its own intellect from uninstructive labor. It would thenceforth progress in civilization much more rapidly than the other. What the mental ability of a race actually achieves is due to the conditions under which it operates, and a race may put on or put off such conditions or have them imposed or removed by other races, for all sorts of reasons.

From all these facts each student may make his own estimate of the original mental differences of races, and learn at least the need of more actual measurements of race differences and of intelligence in interpreting them. My own estimate is that greater differences will be found in the case of the so-called "higher" traits, such as the capacity to associate and to analyze, thinking with parts or elements, and originality, than in the case of the sensory and sensori-motor traits, but that there will still be very great overlapping. Calling the difference between the original capacity of the lowest congenital idiot and that of the average modern European 100, I should expect the average deviation of one pure race* from another in original capacity to be below 10 and above 1, and the difference between the central tendencies of the most and the least gifted races to be below 50 and above 10. I should consider 3 and 25 as reasonable guesses for the two differences.

Even if the differences were far larger than these, the practical precept for education would remain unchanged. It is, of course, that selection by race of original natures to be educated is nowhere nearly as effective as selection of the superior individuals regardless of race. There is much overlapping and the differences in original nature within the same race are, except in extreme cases, many times as great as the differences between races as wholes.

*Defining a pure race arbitrarily as one whose ancestry has less than 1 per cent of community with that of any other race for at least 20 generations back.

CHAPTER XI

The Influence of Immediate Ancestry or Family

There is possibly even less agreement about the amount of influence of immediate ancestry or 'family' than about that of remote ancestry or race. This is the more to be regretted because all the social sciences and especially education need as a starting point precise knowledge of the differences in original mental makeup within the human species and of their relation to immediate ancestry.

The problem naturally resolves itself into two,—the measurement of the resemblance of individuals of like ancestry and the subtraction of a proper allowance for their likeness in training. Or, more exactly, we have to measure the amount by which the likeness of individuals of like ancestry surpasses the likeness of individuals of different ancestry, and subtract from it the amount due to their greater likeness in training than that found in the case of individuals of different ancestry. Measurements of the greater *differences* of *un*related individuals with an allowance for the greater differences in their training would serve the same end. But the effect of differences in ancestry in producing differences in intellect and character is more easily measured by the effect of similarity or identity in ancestry in decreasing such differences. The facts to be considered are then measurements of resemblance and allowance for like training.

THE VARIABILITY OF INDIVIDUALS OF THE SAME SEX AND ANCESTRY

Resemblance, not repetition, is to be measured. To say that a man's original nature depends upon his ancestry does not mean that it is an exact facsimile of any one or any combination of his ancestors. There is no reason to believe that four sons of the same parents and consequently of the same total ancestry will have the same original natures. Indeed we know they will not, save by chance. For twins who have presumably in some cases identical or nearly identical antenatal influences and nurture may vary widely in both physical and mental traits. What ancestry does is to reduce the variability of the offspring and determine the point about which they do vary.

Take, for instance, the capacity to form intelligent habits or associations amongst sense impressions, ideas and acts. The number of associations between situation and act, the number, that is, of things an animal can do in response to the multitude of conditions of life, varies tremendously throughout the animal kingdom. The free swimming protozoa studied by Professor Jennings had in addition to the common physiological functions hardly more than a single habit. The sum of the life of *Paramoecium* is to eat, breathe, digest, form tissues, excrete, reproduce, move along in a steady way, and when passing from certain media into others to stop, back, turn to the aboral side and move along again as before. At the other extreme is a cultivated human being whose toilet, table manners, games, speech, reading, business, etc., involve hundreds of thousands of associative habits.

If now we take a thousand descendants of human beings and count up the number of associative habits displayed by each we shall of course find a great variability. Some of our thousand human offspring will learn fewer things than some dogs and cats. Some of them may learn many more than any of the parents from whom they sprang. But on the whole the offspring of human beings will vary about the human average instead of about the general animal average, and the average deviation of the human group will be far less than that of the whole animal kingdom.

To illustrate again, the children of parents who are, say, 3 inches above the average of the general population in stature will vary not about the general average, but about a point 2 inches above it; and will differ one from another only about ten seventeenths as much as one adult of the general population differs from another.*

Immediate ancestry will then, when influential, cause children to deviate from the general average toward the condition of their parents and to vary less among themselves than would the same number of unrelated individuals.

It might seem at first sight that two individuals of the same sex, race and parentage, two brothers or two sisters, should, if ancestry counted at all, have identical original natures and differ only in as far as different environmental forces affect them. Common observation shows this to be false, but common thinking does not always or often understand that it is false just because immediate ancestry does count. If ancestry did not count, either all men would by original nature be identical, or the variations among them would all be miracles. If ancestry did not count, two brothers might well be identical in original nature, for all human males might be. But if ancestry is a force, it is certainly a variable one, the germs produced by any one parent being somewhat different among themselves for the same reason that the germs produced by all parents together vary still more. If the germs differ at all, the differences are likely to be less amongst the germs of any one human being than amongst an equal number from all men, but the differences are not at all likely to be reduced to zero.

In all thought of inheritance, physical or mental, one should always remember that children spring, not from their parents' bodies and minds, but from *the germs of those parents.* The qualities of the germs of a man are what we should know in order to prophesy directly the traits of his children. One quality these germs surely possess. They are variable. Discarding syntax and elegance for emphasis, we may say that the

*This illustration is based on the data reported by Galton in *Natural Inheritance.*

germs of a six-foot man include some six-feet germs, some six-feet-one germs, some six-feet-two, some five-feet-eleven, some five-feet-ten, etc. Each human being gives to the future, not himself, but a variable group of germs. This hypothesis of the variability of the germs explains the fact that short parents may have tall sons; gifted parents, stupid sons; the same parents, unlike sons. We have to measure the amount of resemblance, not the frequency of identity.

METHODS OF MEASURING RESEMBLANCE

In order to understand the measurements of the resemblance of related individuals which are to be reported in this chapter, it will be helpful to consider one sample case in some detail. Table 33 gives (in columns 1 and 2) the facts concerning the cephalic index (width of head divided by length of head) in the case of 78 individuals, comprising 39 pairs of twins. Columns 3 and 4 give the same facts expressed as deviations from .813, which is the central tendency for cephalic index in the population from which these children were chosen. That is, the first four columns of the first line of the table read,—"Of two twins one had a cephalic index of .788, the other of .790; one was .025 below* the average of the population, the other was .023 below it."

Column 5 gives the difference between the indices of the twins in each of the 39 pairs. Column 6 gives the algebraic product of the deviations of the indices of the twins from the average cephalic index of the population.

Consider first column 5, in connection with this question, "How much smaller is the difference between twin and twin of a pair than that between two children in general?" Column 5 gives the former differences; the latter can be got by getting all the differences between each child and the other seventy-seven. Thus we find individual 1 differing from individuals

*Below meaning more long-narrow-headed, or "dolichocephalic."

TABLE 33.

Resemblances of Twins in Cephalic Index.

Pairs of Twins	1 Cephalic Indices of 1—50. × 1000	2 Cephalic Indices of 101—150. × 1000	3 Deviations of 1—50 from Central Tendency.	4 Deviations of 101—150 from Central Tendency.	5 Differences 1—101, 2—102, etc.	6 Products. Deviation of 1 × deviation of 101, deviation of 2 × deviation of 102, etc.	
1 and 101	788	790	—25	—23	2	575	.....
2 " 102	780	773	—33	—40	7	1320	.....
5 " 105	796	812	—17	—1	16	17	.....
6 " 106	816	826	3	13	10	39	.....
7 " 107	849	837	36	24	8	864	.....
8 " 108	853	859	40	46	6	1840	.....
9 " 109	777	766	—36	—47	11	1692	.....
10 " 110	801	770	—12	—43	31	516	.....
11 " 111	745	738	—68	—75	7	5100	.....
12 " 112	804	773	—9	—40	31	360	.....
13 " 113	823	859	10	46	36	460	.....
16 " 116	765	763	—48	—50	2	2400	.....
17 " 117	778	788	—35	—25	10	875	.....
18 " 118	836	836	23	23	0	529	.....
19 " 119	839	797	26	—16	42	....	—416
20 " 120	761	768	—52	—45	7	2340	.....
21 " 121	786	783	—27	—30	3	810	.....
23 " 123	806	831	—7	18	25	....	—126
24 " 124	778	784	—35	—29	6	1015	.....
25 " 125	826	850	13	37	24	481	.....
26 " 126	802	793	—11	—20	9	220	.....
27 " 127	834	839	21	26	5	546	.....
28 " 128	763	817	—50	4	54	....	—200
29 " 129	827	843	13	30	17	390	.....
30 " 130	814	827	1	14	13	14	.....
31 " 131	789	773	—24	—40	16	960	.....
32 " 132	819	828	6	15	9	90	.....
34 " 134	758	725	—55	—88	33	4840	.....
35 " 135	817	780	4	—33	37	....	—132
36 " 136	786	836	—27	23	50	....	—621
37 " 137	817	830	4	17	13	68	.....
38 " 138	827	846	13	31	18	403	.....
40 " 140	850	793	37	—20	57	....	—740
41 " 141	838	846	25	31	6	775	.....
43 " 143	783	786	—30	—27	3	810	.....
44 " 144	785	826	—28	13	41	....	—364
45 " 145	899	856	86	43	43	3698	.....
47 " 147	832	828	18	14	4	252	.....
50 " 150	836	838	23	25	2	575	.....

101, 2, 102, 5, 105, 6, 106, etc., by 2, 8, 15, 8, 24, 28, 38, 61, 49, etc. Individual 101 differs from 2, 102, 5, 105, 6, 106, etc., by 10, 17, 6, 22, 26, 36, 59, 47, etc. Individual 2 differs from 102, 5, 105, 6, 106, etc., by 7, 16, 32, 36, 46, etc. The differences between twin and twin range from 0 to 57, averaging 18. If the reader should figure out the difference between two children of the same age but not of the same family, he would find it to range from 0 to 174 or more and to average about 40. This is the simplest method of measuring the degree to which individuals of like ancestry resemble one another more than individuals of unlike ancestry do. But it is for many reasons not the most convenient way and is rarely used.

Consider now columns 3 and 4 in connection with the question, "How much oftener are the twins both short-wide-headed or both long-narrow-headed than are two children in general?" Both twins are thus alike (that is, both deviate from the average in the same direction) in 32 out of the 39 pairs, or in 82 per cent of the cases. Two children taken at random will be thus alike in 50 per cent of the cases. *'Greater than 50 per cent frequency of deviation of related individuals in the same direction from the central tendency,'* is another simple measure of their greater resemblance than that of the non-related individuals.

Consider now columns 3 and 4 in connection with this question, "What is the ratio of one twin's deviation from the C. T.* to that of his mate?" We find these 78 ratios, using the ratio of 1's deviation to 101's and also 101's to 1's and so on, to be $\frac{25}{23}$, $\frac{23}{25}$, $\frac{33}{40}$, $\frac{40}{33}$, $\frac{17}{1}$, $-\frac{1}{17}$, $-\frac{3}{13}$, $\frac{13}{3}$, etc., fourteen of them being minus quantities. The central tendency (Median) of these ratios is .87. Now if twin and twin always possessed perfect resemblance—if the two indices of each pair were alike—these ratios would each be 1. If we take the ratio of any person's deviation from the C. T. to that of any other person regardless of blood relationship, there will be as many minus

*C. T. is used here and later as an abbreviation for central tendency.

quantities as plus quantities and the median of the ratios will be 0. The median of the ratios, each of the deviation of one member of a related pair from the central tendency of the whole group to the similar deviation of the other member, is an important measure of resemblance. When it is 0, it means no greater resemblance than one of the group bears to another taken at random. When it is 1.00 it means perfect resemblance or identity. When it is .5 it means a resemblance halfway between perfect identity and that found between any two persons of the group taken at random. This is for many reasons the most serviceable measure of resemblance between individuals, and it, or a measure that may for the purposes of this chapter be regarded as the same as it, will be used in reporting the resemblances found by various students of heredity. The figures, ranging from 0 to 1.00, called Coefficients of Correlation, designated by the symbol *r*, and used to measure resemblance, are to be thought of as expressing the central tendencies (medians) of series of ratios, such as the series of 78 ratios afforded by the deviation-measures of Table 33.

The central tendency of a series of ratios may, however, be calculated in another form, the so-called Pearson coefficient of correlation or resemblance, which is somewhat less easy to describe. Consider column 6 in connection with the question, "What would these products be if the two twins were identical in the case of every pair?" In the first pair the product is 575. If the twins were identical it would be either 625 (25 times 25) or 529 (23 times 23) according as the second twin was to be identical with the first or the first with the second. In the second pair the product is 1320. With perfect resemblance it would be 1089 or 1600. In the third pair it is 17. With perfect resemblance it would be 289 or 1. In the fourth the corresponding products are 39 and 9 or 169. In the fifth pair they are 864 and 1296 or 576.

Consider now the sum of the five products as it is (2815), and as it would be with perfect resemblance (3308 or 2875). Consider what the sum of the products would be with no greater resemblance than that found between any one of the group and any other one taken at random. A plus deviation would then go with a minus deviation as often as with a plus; in the long run as many of the algebraic products would be negative as positive; and their sum would be 0. The proportion which the sum of the products of the related deviations is of what that sum would be with perfect resemblance is thus a measure of the amount of resemblance.

$$\text{This proportion} = \frac{x_1 y_1 + x_2 y_2 + x_3 y_3 \cdots x_n y_n}{\sqrt{(x_1^2 + x_2^2 + x_3^2 \cdots x_n^2)(y_1^2 + y_2^2 + y_2^2 \cdots y_n^2)}}$$

$$\text{or } \frac{\Sigma\, x.\, y}{\sqrt{\Sigma x^2 . \Sigma y^2}}$$

x_1, x_2, x_3, etc. being the deviations of the first members of the related pairs, and y_1, y_2, y_3, etc. being the corresponding deviations of the second members of the related pairs. Σ = 'sum of the series of.'

Thus for the 5 pairs above

$$r = \frac{2815}{\sqrt{3308 . 2875}}$$

When $\frac{\Sigma x.y}{\sqrt{\Sigma x^2 . \Sigma y^2}} = 1$, resemblance is perfect.

When $\frac{\Sigma x.\, y}{\sqrt{\Sigma x^2 .\, \Sigma y^2}} = 0$, there is no greater resemblance than exists between the members of the group taken at random. A resemblance calculated by this, the so-called product-moment method, devised by Karl Pearson, may for our purposes be considered as meaning the same thing as the median of the series of ratios previously described.

If the reader still feels a certain insecurity and unreality about the values of $r = 0$, $r = .1$, $r = .2$, $r = .35$, $r = .75$, $r = .9$, and the like, he can make these coefficients of resemblance or correlation living realities by artificially so pairing the following series as to get from them various values of r, either as r = the median of the $\frac{x}{y} + \frac{y}{x}$ ratios or as $r = \frac{\Sigma x.y}{\sqrt{\Sigma x^2 . \Sigma y^2}}$

Series 1. (x) +1+1+1+1+3+3+3+5+7+11−1−1−1−1 −3−3−3−5−7−11

Series 2. (y) +1+1+1+1+3+3+3+5+7+11−1−1−1−1 −3−3−3−5−7−11

Thus, pairing them in order, $r = 1$, while pairing them (+1+3) (+1−1) (+1+1) (+1−1) (+3+3) (+3+5) (+3−1) (+5+1) (+7+11) (+11+1) (−1+7) (−1−3) (−1−5) (−1+3) (−3+1) (−3−1) (−3−7) (−5−11) (−7−3) (−11−3), $r = .54$. It will be specially useful to pair them so as to get $r = -1$, so as to get $r = 0$, so as to get $r =$ about .2, and so as to get $r =$ about .8.

MEASUREMENTS OF RESEMBLANCE IN RELATED INDIVIDUALS

Before describing the similarities of closely related individuals in mental traits I shall present the results of studies in the case of some physical traits which will prove that heredity is a *vera causa,* since, in them, similarity of training is out of the question as a cause of the similarities found.

The coefficient of correlation between brothers in the color of the eyes is, according to Pearson, .52. But parents could not, if they would, exert any environmental influence upon the color of their children's eyes. The fraternal resemblance must be due to the resemblance in ancestry.

In height Pearson finds the coefficient of correlation between father and son to be .3, and that between brother and brother to be .5. In other words, a son, on the average, deviates from the general trend of the population by .3 the amount of his father's deviation, a brother by .5 the amount of his brother's. Now no one can imagine that tall fathers try especially to make their sons tall. Nor will the class 'men two inches above the average height' feed their children any more than men one inch above it.

The coefficient of fraternal correlation in the case of the cephalic index (ratio of width to length of head) is, according to Pearson, .49. Here it is utterly incredible that fathers do anything to their children that would tend to produce in them similar indices.

Finally take color of hair. Fraternal correlation is, according to Pearson, .55. Here again home influence could not cause one whit of the resemblance.

Immediate ancestry can and does, apart from any other force, cause in whole or in part the abmodality, or deviation from the C. T. of his race, of an individual in the case of stature, cephalic index and eye color. There is no reason to suppose that the brain is less influenced by ancestry than are the tissues that cause height, or the shape of the skull bones that causes cephalic index, or the deposits of pigment that cause eye color. Immediate ancestry is thus a probable cause for original mental nature. And when there is doubt as to the choice between it and the environment as the cause of differences in mental traits of individuals at any age, it must not be forgotten that the influence of the latter is, after all, largely a matter of speculation, while the influence of ancestry is in physical traits a demonstrated fact.

Deafness may be considered a physical trait because it is due to physical causes, but so are all mental traits. The real difference is that we know more about the causes in the one case than in the others. The manifestation and results of deafness are certainly mental traits.

The brother or sister of a person born deaf is found to be deaf in 245 cases out of 1,000, almost one case out of four. The number of deaf persons amongst 1,000 brothers and sisters of hearing individuals is not known exactly. but it is certainly less than 1, probably much less. That is, a person of the same ancestry as a congenitally deaf person is at least 245 (probably many more) times as likely to be deaf as a person of the same ancestry as a hearing person. The child of two parents both of whom were born deaf is at least 259 (probably many more) times as likely to be deaf as the child of two hearing parents [Fay, '98, p. 49]. In this case, as with the physical traits described, there is no reason to impute any efficacy to training. Parents born deaf would take pains to *prevent* deafness in their children.

Mr. E. L. Earle ['03] measured the spelling abilities of some 600 children in the St. Xavier school in New York by careful tests. As the children in this school commonly enter at a very early age, and as the staff and methods of teaching remain very constant, we have in the case of the 180 pairs of brothers and sisters included in the 600 children closely similar school training. Mr. Earle measured the ability of any individual by his deviation from the average for his grade and sex and found the coefficient of correlation between children of the same family to be .50. That is, any individual is on the average 50 per cent as much above or below the average for his age and sex as his brother or sister.

Similarities in home training might theoretically account for this, but any one experienced in teaching will hesitate to attribute much efficacy to such similarities. Bad spellers remain bad spellers though their teachers change. Moreover, Dr. J. M. Rice in his exhaustive study of spelling ability ['97] found little or no relationship between good spelling and any one of the popular methods, and little or none between poor spelling and foreign parentage. Yet the training of a home where the parents do not read or spell the language well must be a home of relatively poor training for spelling. Cornman's more careful study of spelling ['01] supports the view that ability to spell is little influenced by such differences in school or home training as commonly exist.

These facts make it almost certain that immediate ancestry does count somewhat in producing the likenesses and differences found amongst men in mental traits. In the measurements now to be reported, the influence of family training enters as a more probable alternative cause of the resemblance. I shall in each case give the measurement of resemblance made and the allowance for likeness in home training suggested by the author.

The first serious study of the inheritance of mental traits was made in the 60's by Francis Galton and reported in *Hereditary Genius* ['69, '92]. He examined carefully the careers of the relatives of 977 men each of whom would rank as one man in four thousand for eminent intellectual gifts. They had relatives of that degree of eminence as follows:—fathers 89, brothers 114, sons 129, all three together 332; grandfathers 52, grandsons 37, uncles 53, nephews 61, all four together 203. The probable numbers of relatives of that degree of eminence for 977 average men are as follows:—fathers, brothers, and sons together, 1; grandfathers, grandsons, uncles and nephews all together, 3. Galton argues that the training due to the possession of eminent relatives can not have been the cause of this superior chance of eminence in the relatives of gifted literary men and artists.

He says: "To recapitulate: I have endeavored to show in respect to literary and artistic eminence—

1. That men who are gifted with high abilities—even men of class E—easily rise through all the obstacles caused by inferiority of social rank.

2. Countries where there are fewer hindrances than in England, to a poor man rising in life, produce a much larger proportion of persons of culture, but not of what I call eminent men. (England and America are taken as illustration.)

3. Men who are largely aided by social advantages are unable to achieve eminence, unless they are endowed with high natural gifts."

Galton demonstrates that the adopted sons of popes do not approach equality in eminence with the real sons of gifted men. He so orders his studies of men eminent in other fields as to leave very slight basis for one who argues that training and opportunity rather than birth caused the eminence attained. Finally, Galton's own opinion, that of an eminently fair scientific man based upon an extensive study of individual biographies, may safely be taken with a very slight discount. He says:—"I feel convinced that no man can achieve a very high reputation without being gifted with very high abilities."

The historic importance of Galton's *Hereditary Genius,* the originality and ingenuity of its author and its substantial results, should make his book the first to be read by every student of mental inheritance.

Loewenfeld ['03] confirms Galton's estimate of the resemblance in intellect and energy amongst descendants from similar near ancestry.

In 1889 Galton published his *Natural Inheritance,* the results of more precise studies* of resemblances amongst related individuals in stature, eye color, the artistic faculty and diseases. He found the resemblance between parents and their children in the mental trait studied (artistic faculty) to be a little greater than in the case of stature. The essential facts from which this inference is drawn are that in 30 families where both parents were artistic 64 per cent of the children were so, whereas in 150 families where neither parent was artistic only 21 per cent of the children were so ['89, p. 218]. No attempt is made to divide the causation of this resemblance between birth and training.

Pearson ['04] secured ratings by teachers of about 2,000 pairs of siblings with respect to the following traits:—ability, vivacity, conscientiousness, popularity, temper, introspection or self-consciousness, assertiveness and handwriting. In ability and handwriting the measure was a grade from one to six according to ability; in temper from one to three. In the other traits the individual was put into an upper or a lower class. Such material is not well suited for measuring resemblance, chiefly for two reasons. First, the measurements are very coarsely made, and so have a large chance inaccuracy. A slight prejudice or ignorance in the teacher making the measurement may put a boy who is really above the average in one of these traits below it. If errors of judgment displace a boy on the scale at all, they will alter his position by half the total scale in five of these traits, by a third of the scale in one and by a

*These studies were reported in various memoirs from 1871 to 1887.

sixth in the other. Such chance errors are sure to occur. Their effect is to make the resemblances obtained from the teachers' ratings *lower* than the real resemblances. Spearman ['04a, p. 97ff.] has estimated that if these teachers' ratings suffer from such chance errors as he himself finds in such ratings, the obtained resemblances (which are about .5) are very much too low, the real resemblances required to produce a measurement of $r = .5$ when 'attenuated' by the teachers' chance errors, being over .8. Pearson denies that the teachers' ratings were thus inaccurate, claiming that they displaced a pupil from his proper class only rarely. Spearman possibly overestimates the chance errors of a teacher's judgment of the intellect or character of a boy well known to him, but Pearson certainly underestimates them, when he makes no allowance at all for them.

He measured them only in the case of 'ability,' which would be the easiest of these traits for a teacher to grade accurately, and which was graded by a six-step scale. Three teachers in a school graded the same pupils for him* without consultation among themselves, with the result, he writes, that 'the agreement in classification was complete in more than 80 per cent of cases and only differed by as much as two classes in about 5 per cent of cases.' ['04, p. 161.] But such gradings are not really independent, if the pupils are talked about by their teachers.

The second error acts in the reverse direction. Any teacher has his own idea of 'quiet,' noisy,' 'keen' in conscientiousness or 'dull' in conscientiousness, and the like. Whether he shall rate a boy as 'self-conscious' or 'unself-conscious' depends only in part on what the boy's nature is. In part it depends on where the teacher draws the line between being self-conscious and being unself-conscious. Suppose, for instance, John and James to be really at points *Jo* and *Ja* on the scale for self-consciousness represented in Fig. 48, and suppose the average self-

*This being done for some 150 boys and girls in all, distributed in six schools.

consciousness of English boys to be really at point *Av.* on this scale. Consider now ten teachers each of whom is asked, "Is John self-conscious or unself-conscious?" Each of them is in total ignorance of the position of the dividing point, every point above which should be rated 'self-conscious' and every point

FIG. 48. The real unlikeness of Jo. and Ja.

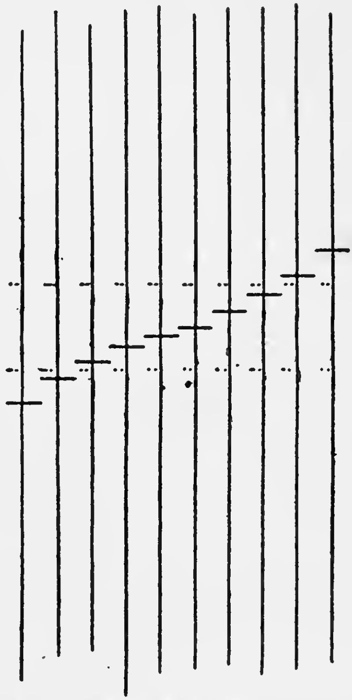


FIG. 49. The specious likeness of Jo. and Ja. when they are rated by teachers who place the dividing-line between self-conscious and unself-conscious too low or too high.

below which should be rated 'unself-conscious.' Each of them makes his own dividing point. These divisions will vary. How much they will vary is not known. Fig. 49 gives an arbitrary estimate of ten such dividing lines (the heavy cross lines) as an illustration. Now John and James, when rated by these ten subjective scales, seem alike in four cases out of ten, though they are really not in this sense alike, one being really above and the other below the real dividing line. Now, in general, any difference in judges in making the subjective divisions of the scales whereby they rate the boys will lead the judges to rate boys who are really different as alike oftener than it will lead them to rate boys who are really alike as different. For there are many more boys near the dividing line than at the extremes, and the brothers of a pair are usually rated by the same judge. Hence the resemblances as calculated from these teachers' ratings are, in so far forth, greater than the real resemblances. How much greater is not known, since the amount of difference between the subjective and the real dividing point is not known. The error would probably be greater in the case of conscientiousness, self-consciousness and shyness, where subjective standards probably vary very much, than in the case of intelligence or popularity.

The average resemblance of .52 obtained by Pearson would thus be raised if one error had been avoided, and lowered if a second error had been avoided. I judge that roughly on the whole the two errors would somewhat nearly balance. Just what would happen to each of the particular resemblances if accurate original measurements had been secured, cannot be more than guessed. Pearson's obtained measures were ['04, p. 155]:—

	Brothers	Sisters	Brother and Sister
Vivacity	.47	.43	.49
Self-assertiveness	.53	.44	.52
Introspection	.59	.47	.63
Popularity	.50	.57	.49
Conscientiousness	.59	.64	.63
Temper	.51	.49	.51
Ability	.46	.47	.44
Handwriting	.53	.56	.48
Average	.52	.51	.52

A third criticism of Professor Pearson's measurements might be that the teachers who reported cases of brothers tended to select brothers whose likeness was notable and to grade them alike oftener than the facts warranted. It is true that if you seek information from people without special training in the field of science concerned, the answers obtained are well-nigh sure to be compounds of fact and prejudice, but there is apparently no reason to believe that the teachers had, apart from the directions sent to them, any prejudicial belief in fraternal similarity.

In two ways the directions for making the judgments may have tended to suggest the existence of resemblances. The directions read:—"1. The object of this investigation is twofold: (1) To ascertain the degree of resemblance, mental and physical, between children of the same parents.............. 2. The measurements and estimates are to be made on (I) Pairs of brothers (white data paper). (II) Pairs of sisters (pink data paper). (III) Pairs of brothers and sisters (blue data paper)." ['04, p. 161.] It would have been far better to have had all the children in the school rated each on a sheet by himself and to have said nothing about the purpose for which the measurements were to be used.

There is one more fact to be noted which must lessen confidence in these measurements still further. Using just the same method of securing the data as in the cases discussed, Professor Pearson gets, from ratings of *athletic* or *non-athletic,* resemblances of .72 for brothers, .75 for sisters and .49 for brother and sister ['04, p. 154]. Now we know that if these measurements are free from error in all respects save the mere *chance* errors of judgment of the teachers, the real resemblances must be even higher than .72, .75 and .49. But so close resemblance as a result of original nature alone is absurd. There is nothing like it elsewhere in fraternal resemblances, save in twins. In the case of athletic power Professor Pearson's figures surely measure the differences of schools in devotion to athletics or the differences of the judges as to where the dividing line is between *athletic* and *not athletic,* as well as the resemblances of brothers and sisters in original nature.

Professor Pearson thinks that no allowance is due for the similarities of home training because the resemblances in health.

eye-color, hair-color and five other physical traits average .54, .53 and .51 for the three groups of pairs. The resemblances in these physical traits cannot be due to the fact that the brothers have the same home environment, he says, yet the resemblances are as great as in mental traits. Professor Pearson lays great stress on the exact equality of physical and mental resemblance, but since his measures of the latter are subject to two important sources of error, this exact correspondence is of doubtful significance.

Unless one is a blind devotee to the irrepressibility and unmodifiability of original mental nature, one cannot be contented with the hypothesis that a boy's conscientiousness or self-consciousness is absolutely uninfluenced by the family training given to him. Of intelligence in the sense of ability to get knowledge, rather than amount of knowledge got, this might be maintained. But to prove that conscientiousness is irrespective of training is to prove too much. One fears that Professor Pearson may next produce coefficients of correlation to show that the political party a man joins, the place where he lives, and the dialect he speaks are matters of pure inheritance uninfluenced by family training

The reader may at this stage be in some doubt as to precisely what Professor Pearson's measurements give as a probable similarity of brothers' original natures. I share this doubt, and from estimating the different sources of error can do no more than expect that adequate mental measurements with the effects of similarity of training eliminated would leave resemblances of from .3 to .5.

Dr. Frederick Adams Woods has reported in *Mental and Moral Heredity in Royalty* ['06], a work which appeared first in the Popular Science Monthly in 1902 and 1903, measurements of the resemblances in intellect and in morals of many individuals chosen from the royal families of Europe. Dr. Woods gave to each person of the 671 studied a rating from 1 to 10 on a scale for intellect—1 representing feeble-mindedness or imbecility; 10 such gifts as those of William the Silent, Frederick the Great and Gustavus Adolphus; and 2, 3, 4, 5, etc., steps at equal intervals between, in his opinion. These ratings represented Dr. Woods' impressions from reading the statements of historians and biographers about these individuals. He gave similar ratings for morality.

The ratings assigned by Dr. Woods are, of course, not accurate. No one man's ratings of nearly seven hundred historical personages could be. The effect of this chance inaccuracy would be to make all his measurements of resemblance lower than the real resemblance. He may also have erred from an unconscious prejudice by rating as too much alike individuals who were closely related. This error would tend obviously to make his estimate of resemblance too high. His ratings are given in full and so far nobody has proved or even suggested that they are thus biased.

There is still another chance for error. The reputation of a prince may be a peculiarly unfair measure of his ability. A son whose gifted father has brought the nation's affairs into a prosperous condition may thereby get, in histories and biographies, an unduly high rating; whereas a son who must strive against the unfavorable conditions produced by a stupid father, may thereby incur an undeserved repute of inefficiency. This is, however, no more plausible a supposition than the opposite one that a moderately gifted son would be rated too low by contrast with a gifted father and too high if his predecessor had been a marked failure. On the whole Dr. Woods' ratings seem little subject to error other than chance inaccuracy, so that the resemblances calculated from them are probably too low rather than too high.

The bulk of this work is devoted to the concrete description of particular stocks and the consequences of particular matings. To Dr. Woods at least the likenesses and differences of these men and women seem due in very large measure to their likeness and difference in ancestry, not only in the case of the degree of intellect and of morals but also in the direction of interests and in minor features of temperament. The general results of the measurements of resemblance are as follows:—

"By taking the records of each country separately and analyzing them minutely, we have seen how almost perfectly established heredity appears to be as a cause of decided mental and moral peculiarities, wherever found. Instead of treating each country separately, the entire number of interrelated persons will now be studied as if they were arranged on a single chart, according to blood relationship. If such a great chart were constructed, we should see the geniuses, or (9) and (10) grades, not scattered at random over its entire surface, but isolated little groups of (9) and (10) characters (the individuals within each group contiguous to each other) would be found here and there. One such group would be seen centering around Frederick the Great, another around Queen Isabella, of Spain, another in the neighborhood of William the Silent, and still a fourth with Gustavus Adolphus as a center. These would constitute the largest groups of closely related (9) and (10) characters. There would also be a few other groups of two or three geniuses each.

"Those in the lowest grades for intellect would also be found close to others of the lowest type, and would fall especially in Spain and Russia, in which countries we have seen an inherited insanity. There would be certain regions composed almost entirely of grades from (4) to (7). These would cover the greater part of the chart and include the houses of Hanover, Saxe-Coburg-Gotha, Reuss, Mecklenburg, Hapsburg in Austria, Holstein, Denmark, Saxony, Savoy, Orleans and modern Portugal." ['06, pp. 265-266.]

The general tendency to resemblance he finds to be:—

In intellect:—

Offspring and fathers, $r = .30$;

Offspring and grandfathers, $r = .16$;

Offspring and greatgrandfathers, $r = .15$

In morals:—

Offspring and fathers, $r = .30$

Offspring and grandfathers, $r = .175$

In the case of intellect there is a peculiarity in the resemblance. Men of grades 1 to 4 have equally gifted fathers; only from then on is there a rise in paternal gifts in proportion to the gifts of the offspring. In moral qualities it is not so.

Dr. Woods thinks that little or no allowance need be made for greater similarity of environment for son and father or grandfather than existed for sons of royal families in general. He says that, while educational opportunities have been unequal, the "advantages and hindrances must have always been of an accidental character, depending on various causes, and their distribution would occur largely at haphazard throughout the entire number of collected persons (832); and could not account for the great group of mediocrity and inferiority, like the houses of Hanover, Denmark, Mecklenburg, and latter Spain, Portugal and France." ['06, p. 284.]

So also the advantages of high military or political office have been, in his opinion, distributed "at random throughout the entire number and could not produce the grouping by close blood relationship found throughout this entire study." ['06, p. 285.]

He tests one environmental influence by the facts, namely, the advantage of succession to the throne.

"There is one peculiar way in which a little more than half of all the males have had a considerable advantage over the others in gaining distinction as important historical characters. The eldest sons, or if not the eldest, those sons to whom the succession has devolved, have undoubtedly had greater opportunities to become illustrious than those to whom the succession did not fall by right of primogeniture. I think every one must feel that perhaps much of the greatness of Frederick II, of Prussia, Gustavus Adolphus, and William the Silent, was due to their official position; but an actual mathematical count is entirely opposed to this view. The inheritors of the succession are no more plentiful in the higher grades than in the lower. The figures below show the number in each grade who came into power by inheriting the throne.

Grades	1	2	3	4	5	6	7	8	9	10
Total number in each Grade	7	21	41	49	71	70	68	43	18	7
Succession Inheritors	5	14	26	31	49	38	45	23	12	4
Per cent	71	67	63	64	69	54	67	54	67	57

It is thus seen that from 54 to 71 per cent inherited the succession in the different grades. The upper grades are in no way composed of men whose opportunities were enhanced by virtue of this high position. Thus we see that a certain very decided difference in outward circumstances—namely, the right of succession—can be proved to have no effect on intellectual distinction, or at least so small as to be unmeasurable without much greater data. The younger sons have made neither a poorer nor a better showing." ['06, pp. 285-286.]

His conclusion is:—"The upshot of it all is, that as regards intellectual life, environment is a totally inadequate explanation. If it explains certain characters in certain instances, it always fails to explain as many more; while heredity not only explains all (or at least 90 per cent) of the intellectual side of character in practically every instance, but does so best when questions of environment are left out of the discussion. Therefore, it would seem that we are forced to the conclusion that all these rough differences in intellectual activity which are susceptible of grading on a scale of ten are due to predetermined differences in the primary germ-cells." ['06, p. 286.]

In the case of the resemblances in morality Dr. Woods is less emphatic in denying that any discount for similarity in training is required. In fact his attitude is not clear. He says:

"The conclusion seems to be, therefore, that even in the moral side of character, inherited tendencies outweigh the effects of surroundings, for the reason that, applied to all the characters, heredity is able to explain almost every one,—there being but a slight error from the expected,—while environment will only explain a relatively smaller number. I think we can conclude from this that in each individual, inheritance plays, in the formation of morality, a force greater than 50 per cent. Other considerations enable us to go even farther than this. The comparison between maternal and paternal grandsires is significant. Offspring resemble their maternal grandfathers as much as their paternal. Here we test the resemblances under diverse conditions of environment, the conditions of heredity remaining the same, yet we find no weakening of the latter force. Such a result is surprising, for it does seem improbable that environment has *no* influence in the determination of temperament, behavior, and virtue in general; and there is, of course, an ingrained popular belief that it has." ['06, p. 294.]

But only a page later he adds, "All I can say is, that I have made several tests to find a measurable influence of environment apart from inheritance, and have failed to find it in this research." ['06, p. 295.] In discussing the fact that the same stock shows both very bad and very good men and women, he says, "It is these strong contrasts, more than anything else, that must lead us to the conclusion that what we have in Plate 2 [a chart showing moral resemblance of offspring and parents, not reprinted here] is truly the effect of blood relationship, for environment should not cause this distribution. Spain, France, and Russia give us most of the degenerates. In these countries the individuals are closely associated in blood with insanity, epilepsy, or other psychoses. This is itself a coincidence to be explained by those who doubt that morality is much the result of inheritance. ....When strong contrasts are found among the children, we always find strong contrasts among the ancestors." ['06, pp. 290-292.]

In 1905 the author published a report [Thorndike, '05] of measurements of the resemblances of fifty pairs of twins in marking A's on a printed page of capital letters (A test), marking words containing certain combinations of letters (a-t and r-e), marking misspelled words on a sheet containing 100 words (misspelled word test), addition, multiplication and writing the opposites of a set of words. I quote or summarize the essential facts.

The resemblances of twins, resemblance meaning any greater likeness than would be found in a pair of children of the same age and sex picked at random from the school population of New York City, are:—

In the A test	R = .69
In the a-t and r-e tests	R = .71
In the misspelled word test	R = .80+
In addition	R = .75
In multiplication	R = .84
In the opposites test	R = .90

There were two possible sources of error in the measurements:

—namely, (1) the possibly unfair selection of twins and (2) the fact that the two members of a pair were commonly tested together. The method of discovering twins was as follows:—Teachers in certain schools were asked to inquire of their pupils whether any one had a twin brother or sister. All twins so reported were tested. But also frequently some teachers would report that in such and such a school there was a pair of twins. These could then be found quickly and measured. These reported cases were perhaps likely to have been noticed in the first case because of their likeness and so to be an unfair selection. Again, in the New York schools it is usual to separate the sexes after three or four years of school life, and it is a frequent practice to separate them from the start. Twins of like sex are therefore more conveniently obtained and so more often tested than their general frequency would recommend. The amount of the resulting constant error is not, however, great.* The tests were all made by the same individual and in the same way except for unconscious changes. However, in respect to time of day, conditions of weather and light, and such conditions as are determined by family life, *e. g.*, the lack of breakfast, fatigue from a party the previous night, and the like, two twins would, when measured at the same time, be influenced alike. Thus the obtained resemblance would be too large. I can evaluate the amount of the resultant constant error only from general considerations. I believe it to be small. This constant error would also influence the correction made from attenuation, but here would make the obtained resemblance too small. An allowance was made for these two sources of error to the best of the author's ability.

If now these resemblances are due to the fact that the two members of any twin pair are treated alike at home, have the same parental models, attend the same school and are subject in general to closely similar environmental conditions, then (1) twins should, up to the age of leaving home, grow more and

*Of the fifty pairs of twins measured there were three more of the same sex than would be expected in a random selection.

more alike, and in our measurements the twins 13 and 14 years old should be much more alike than those 9 and 10 years old. Again, (2) if similarity in training is the cause of similarity in mental traits, ordinary fraternal pairs not over four or five years apart in age should show a resemblance somewhat nearly as great as twin pairs, for the home and school conditions of a pair of the former will not be much less similar than those of a pair of the latter. Again, (3) if training is the cause, twins should show greater resemblance in the case of traits much subject to training, such as ability in addition or in multiplication, than in traits less subject to training, such as quickness in marking off the A's on a sheet of printed capitals, or in writing the opposites of words.

On the other hand, (1) the nearer the resemblance of young twins comes to equaling that of old, (2) the greater the superiority of twin resemblance to ordinary fraternal resemblance is, and (3) the nearer twin resemblance in relatively untrained capacities comes to equaling that in capacities at which the home and school direct their attention, the more must the resemblances found be attributed to inborn traits.

The older twins show no closer resemblance than the younger twins, and the chances are surely four to one that with an infinite number of twins tested the 12-14-year-olds would not show a resemblance .15 greater than the 9-11-year-olds. The facts are:—

The Resemblances of Young and Old Twins Compared

	Twins 9-11	Twins 12-14
1) A test	66	73
2) a-t and r-e tests	81	62
3) Misspelled word test	76	74
4) Addition	90	54
5) Multiplication	91	69
6) Opposites	96	88
Averages	83	70

I have measured the resemblances between siblings (children of the same parents) a few years apart in age only imperfectly, and only in the A test, a-t test and opposites tests. The resemblances are between .3 and .4, or less than half the resemblance found for twins.

The variations in the closeness of resemblance of the twins in the different traits show little, and possibly no, direct correlation with the amount of opportunity for environmental influences. The traits most subject to training (addition and multiplication) do show closer resemblances than the traits least subject to training (the A, a-t and r-e test); but on the other hand show less close resemblances than the traits moderately subject to training (the misspelled word test and opposites test).

The facts then are easily, simply and completely explained by one simple hypothesis: namely, that the natures of the germ cells—the conditions of conception—cause whatever similarities and differences exist in the original natures of men, that these conditions influence body and mind equally, and that in life the differences in modification of body and mind produced by such differences as obtain between the environments of present-day New York City public school children are slight.

We must be careful, however, not to confuse two totally different things: (1) the power of the environment,—for instance, of schools, laws, books and social ideals,—to produce differences in the relative achievements of men, and (2) the power of the environment to produce differences in absolute achievement. It has been shown that the relative differences in certain mental traits which were found in these one hundred children are due almost entirely to differences in ancestry, not in training; but this does not in the least deny that better methods of training might improve all their achievements fifty per cent or that the absence of training, say, in spelling and arithmetic, might decrease the corresponding achievements to zero.

The argument is limited entirely to the causes which make one person differ from another in mental achievements *under the same general conditions of life at the beginning of the* *twentieth century in New York City as pupils in its school system.* If the resemblance of twins has been measured in the case of a group made up partly of New York City school children and partly of children of equal capacity brought up in the wilds of Africa, the variability of the group in addition and multiplication would have increased and the correlation coefficients would rise. They would then measure the influence of original nature plus the now much increased influence of the environment.

Heymans and Wiersma ['06 and '07] sought to measure the influence of heredity by having the parents and children in a family rated by some one who knew them well. Such reports were obtained from physicians through a questionnaire. The ratings were under ninety rubrics of which the following are samples:—

1. Is the person in question active (gesticulating, jumping up from the chair, going up and down the room) or passive and quiet?

2. Is the person in question, in his professional, business, school or home life, always zealous at his work, or only sometimes zealous, or outright lazy?

The resulting ratings suffer from the constant error toward too great resemblance due to the fact that any one family is judged with reference to the same dividing line for a trait, the line being in a different place along the scale for each judge. This source of error is even more mischievous here than in Pearson's study, since here apparently, one judge commonly rated only one family. They also suffer from the random inaccuracy of the judgments which, as was shown on page 237, would make the obtained resemblances too low.

These authors do not measure the resemblances in the usual way nor give data permitting anyone else to do so. In view of the insecurity of their original data it seems best not to enter upon an explanation of their somewhat awkward method of measuring the force of heredity, and not to repeat the figures which are got by this method. Also they do not attempt to estimate an allowance for the influence of similarity in home training, though they state that some such allowance must be made.

They point out, however, that there is reason to believe that the influence of heredity far outweighs the influence of home training, since the resemblance is hardly any greater in traits much subject to the latter influence than in traits little subject to it. They instance industry, sympathy and patriotism on the one hand and emotionality, consolability, and memory on the other.

There exist many facts concerning the ancestry of individuals who themselves are roughly measured as insane, feeble-minded, convicted of crime and the like. So also there exist facts concerning the descendants from individuals so classed. The facts range in definiteness from the proved record of a commitment to an asylum or jail to the opinions of one or more rather untrustworthy persons concerning the nature and career of the individual in question.

These facts are adequate to convince one that immediate ancestry is potent in causing individual differences in intellect and character, but are not adequate to give any precise measure of the amount of its potency. The probable error in rating a man as 'feeble-minded' or 'normal,' 'healthy-minded' or 'neurotic,' 'criminal' or 'law-abiding,' is very great when the rating is made indirectly from the testimony of relatives or acquaintances, and still greater when also the rating is for an individual long dead. Moreover, the investigator is liable to misleading from the very potent suggestions of the facts known concerning other relatives, and from his own beliefs concerning the influence of ancestry. Having found one clearly feeble-minded child of a certain mating of two feeble-minded, his questions concerning other children of that mating are likely to differ from the questions that he would have asked had he been questioning about a child of two gifted people, brother to a clearly gifted child. Even his weighting of the answers to his questions is likely to vary.

I shall not therefore attempt the almost impossible task of estimating the amounts of resemblance found and the amounts of discount for similarities in training to be made, in the case of these studies of ancestry and progeny. I shall only quote representative results obtained in two of the latest and best of such studies, that of Goddard and Kite [Goddard, '12] on mental deficiency and that of Rosanoff and Orr ['11] on insanity.

In connection with the Vineland Institution for the care and education of mentally deficient children, investigations into the antecedents of these children have been made. The inquiry in the case of Deborah K. led back in five generations to a man of good ability who had an illegitimate son Martin by a girl recorded as feeble-minded. Martin married a woman recorded as normal and had nine children, five recorded as feeble-minded, two as normal, one unknown and one who died in infancy. The five feeble-minded children, contracting marriage or illicit unions with individuals of low intellectual power, gave rise to a brood of incompetents who in turn often bred with others like themselves. The result in the great-grandchildren of the five, and in the offspring of these, is a horrid array of human incompetents. From this illegitmate son Martin "have come four hundred and eighty descendants. One hundred and forty-three of these, we have conclusive proof were or are feeble-minded, while only forty-six have been found normal. The rest are unknown or doubtful. Among these four hundred and eighty descendants thirty-six have been illegitimate. There have been thirty-three sexually immoral persons, mostly prostitutes. There have been twenty-four confirmed alcoholics. There have been three epileptics. Eighty-two died in infancy. Three were criminal. Eight kept houses of ill-fame." [Goddard, '12, p. 18 f.]

Rosanoff and Orr ('11) investigated seventy-two families, members from which were inmates in a certain state hospital for the insane. Samples of the results are shown in Figs. 50 and 51.

According to their ratings, the seventeen matings where both parents were neuropathic resulted in 75 offspring of whom 54 were neuropathic, 11 died in infancy, 1 had asthma, 1 was "somewhat odd and possibly abnormal in make-up," while 8 were from 8 to 22 years of age without having shown

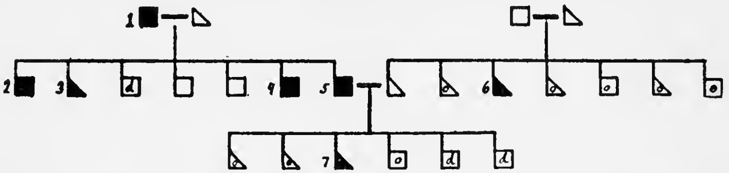


FIG. 50. The Family History of a Case of Manic-depressive Insanity. A square = a male; a triangle = a female; black = a neuropathic individual in the opinion of the investigators; a *d* means "died in childhood;" an *o* means had no offspring. Evidence given that the individual in question was neuropathic is as follows: "1. Insane before death; 2, 'Nervous prostration,' in sanitarium four weeks, recovered; 3, 4, Manic-depressive insanity, in State hospital; 6, Epilepsy, in State hospital." [Rosanoff and Orr, '11, p. 230]

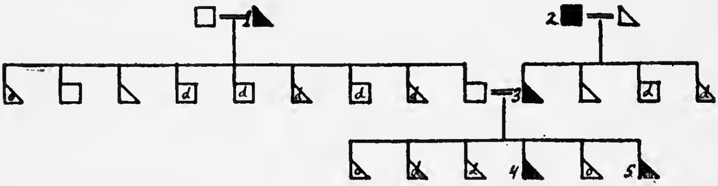


FIG. 51. The Family History of Another Case. The meaning of the symbols is as in FIG. 50. The evidence given is as follows: "1, hysterical when girl, had idea some one was trying to poison her; 2 and 3, Epilepsy; 4, manic-depressive insanity, in State hospital; 5, Very nervous." [Rosanoff and Orr, '11, p. 231]

evidence of neuropathic tendencies. The percentage of neuropathic offspring amongst those living was then 86 or more. The 107 matings where one parent only was neuropathic resulted in 40% of neuropathic cases among the living offspring about whom data were available. The 82 matings where neither parent was rated by them as neuropathic resulted in 27 per cent of neuropathic cases among the offspring living and traced.*

*The reader will understand that these matings, found in 72 families selected in the first instance by the presence of insanity, are not at all a fair sampling of non-neuropathic matings in general. Such matings in general would presumably give a far lower percentage of neuropathic offspring.

THE SPECIALIZATION OF THE INFLUENCE OF NEAR ANCESTRY

In the facts so far given in this chapter, the emphasis has been upon the amount of influence of near ancestry. The degree to which it is specialized is also of importance. How far, for example, do particular talents exist in a man's original nature as a result of his ancestry? Is a man from the beginning organized to be a novelist, or only to be a writer of fiction, or only to be an artist of some sort, or perhaps even only to be a man of ability? How far is this original specialization, if it exists, due to ancestry? It is a plausible statement that individual minds are dependent on heredity only in their rough outlines, the currents of mental activity being fixed only in their general directions and left to take what particular channels circumstances may decide. But this or any other opposite statement must be put in terms of 'how much,' 'how far,' 'in what cases,' to be theoretically satisfying or practically useful.

Galton, who had this problem clearly in mind, notes a number of relevant facts, some of which I quote. Concerning the judges of England between 1660 and 1865 he says:

"Do the judges often have sons who succeed in the same career, where success would have been impossible if they had not been gifted with the special qualities of their fathers? . . .

"Out of the 286 judges, more than *one in every nine* of them have been either father, son or brother to another judge, and the other high legal relationships have been even more numerous. There cannot, then, remain a doubt but that the peculiar type of ability that is necessary to a judge is often transmitted by descent." ['69, '92, pp. 61 and 62.]

Concerning the eminent relatives of eminent statesmen he says:

"Thirdly, the statesman's type of ability is largely transmitted or inherited. It would be tedious to count the instances in favor. Those to the contrary are Disraeli, Sir P. Francis (who was hardly a statesman, but rather a bitter controversialist) and Horner. In all the other 35 or 36 cases in my appendix, one or more statesmen will be found among their eminent relations. In other words, the combination of high intellectual gifts, tact in dealing with men, power of expression in debate, and ability to endure exceedingly hard work, is hereditary." ['69, '92, pp. 103 and 104.]

Similar specialization of inheritance is shown to be the case with the relatives of great commanders, literary men, poets and divines. With men of science the fact is much more pronounced, twenty-two out of the twenty-six eminent sons of eminent scientific men having been eminent in science. This extreme specialization of resemblance is in part due, Galton thinks, to training.

The eminent relatives of eminent painters seem to be well-nigh universally gifted in the same special line. In Galton's list all the relatives mentioned are painters save four. These were gifted in sculpture (2), music (1) and embroidery (1). Finally, "Mendelssohn and Meyerbeer are the only musicians in my list whose eminent kinsmen have achieved their success in other careers than that of music." ['69, '92, p. 231.]

Of course there is, in the case of all of Galton's facts, the possibility that home surroundings decided the special direction which genius took, that really original nature is organized only along broad lines. Moreover, it is difficult to see just what in the nervous system could correspond to a specialized original capacity, say, to be a judge. Still the latter matter is a question of fact, and of the former issue Galton's studies make him the best judge. We should note also that it is precisely in the traits the least amenable to environmental influence, such as musical ability, that the specialization of family resemblance is most marked.

Ellis notes 'a clearly visible tendency for certain kinds of ability to fall into certain [family] groups.' ['04, p. 83.] In his group of 1030 eminent men he finds that:

"Men of letters are yielded by every class, perhaps especially by the clergy, but Shakespeare, and it is probable Milton, belonged to the families of yeomen. The sons of lawyers, one notes, even to a greater extent than the eminent men of "upper class" birth, eventually find themselves in the House of Lords, and not always as lawyers. The two groups of Army and

Medicine are numerically close together, but in other respects very unlike. The sons of army men form a very brilliant and versatile group, and include a large proportion of great soldiers; the sons of doctors do not show a single eminent doctor, and if it were not for the presence of two men of the very first rank—Darwin and Landor—they would constitute a comparatively mediocre group.

"Painters and sculptors constitute a group which appears to be of very distinct interest from the point of view of occupational heredity. In social origin, it may be noted, the group differs strikingly in constitution from the general body, in which the upper class is almost or quite predominant. Of 63 painters and sculptors of definitely known origin, only two can be placed in the aristocratic division. Of the remainder 7 are the sons of artists, 22 the sons of craftsmen, leaving only 32 for all other occupations, which are mainly of lower middle class character, and in many cases trades that are very closely allied to crafts. Even, however, when we omit the trades as well as the cases in which the fathers were artists, we find a very notable predominance of craftsmen in the parentage of painters, to such an extent indeed that while craftsmen only constitute 9.2 per cent among the fathers of our eminent persons generally, they constitute nearly 35 per cent among the fathers of the painters and sculptors. It is difficult to avoid the conclusion that there is a real connection between the father's aptitude for craftsmanship and the son's aptitude for art.

"To suppose that environment adequately accounts for this relationship is an inadmissible theory. The association between the craft of builder, carpenter, tanner, jeweller, watchmaker, woodcarver, rope maker, etc., and the painter's art is small at the best, and in most cases non-existent." ['04, pp. 84, 85.]

Ellis adds, "It may be noted that Arreat (*Psychologie du Peintre,* 1892, ch. 11) in investigating the heredity of 200 eminent European painters, reached results that are closely similar to those I have reached in my smaller purely British group. He found that very few were of upper class social rank, and those not usually among the most important, while nearly two-thirds of the whole number were found to be the sons either of painters or of workers in some art or craft. He refers to the special frequency of jewellers among the fathers. I may remark that in my list, working jewellers and watch-makers occurred twice, a small number, but relatively large considering that there are only three fathers of this occupation in the total parentage of British men of ability." ['04, p. 85.]

Pearson's measurements concern somewhat specialized traits from the start. In so far as his results are trustworthy they show equal resemblance in so specialized a trait as temper or handwriting with that found for 'ability.' Woods notes the inheritance of literary ability, of common sense, of insanity and other somewhat specialized traits in the royal families. In the twins whom I measured the resemblance is almost or quite as great in the case of any one of the six tests (see p. 247) as in the average ability in all of them.

From the work of Burris ['03], if his results are properly corrected for the chance inaccuracies in the original measures, it would appear that the ability to do well in some one high school study is nearly or quite as much due to ancestry as is the ability to do well in the course as a whole.

What knowledge we have thus supports the view that a man's original nature is organized by inheritance in great detail, particular traits and complexes of traits showing similarity between father and son or brother and brother.

Mental inheritance is specialized also in the further sense that two individuals alike in one trait as a result of heredity need not be equally alike in some other trait, though the latter be in general equally subject to inheritance. For example, a pair of twins may be indistinguishable in eye color and stature but notably different in hair color and in tests of intellect.

To measure the extent of this specialization of resemblance, exact measures of the resemblance of individual relatives are heeded. The procedure required to measure accurately the resemblance between one individual and another in one trait is somewhat intricate, and I shall not describe it here. The reader may verify the accuracy of the measures which I give by examining *Measurements of Twins,* [Thorndike, '05] sections 16 to 18 and 21 to 23. Here I shall simply give the resemblances in efficiency in the A test, a-t test, and misspelled word test of the ten pairs of twins who showed the closest resemblance in efficiency in addition, multiplication, and writing opposites. These show (Table 34) that very close resemblance in efficiency in 'association' may go with actual antagonism, or greater unlikeness than that shown by children of the same age, taken at random, in efficiency in tests of 'perception.'

TABLE 34.

SPECIALIZATION OF RESEMBLANCE IN TWINS.

Twin pair		Resemblance in association: resemblance in efficiency in add., mult. and oppos. combined	Resemblance in perception: resemblance in A test, word test and misspelled word test combined.
2	102	1.00—	.28
17	117	.99	.83
20	120	1.00—	—.88
21	121	.99	—.86
26	126	.95	1.00—
27	127	.94—	1.00—
33	133	.95	1.00—
36	136	.94	—.73
46	146	1.00—	—.98
48	148	1.00—	.97

In the *Measurements of Twins* I have also shown in detail that twins may be indistinguishable in any one physical trait without being at all similar in certain others, or in certain mental traits. For example, a pair of twins may show resemblance of .95 to 1.00 in bodily measurements other than head measurements and of 0 in head measurements.

As a consequence of the specialization of resemblance the fact of 'a child resembling his parents' is seen to be in reality 'the mental traits of a child resembling each the corresponding mental trait in his parents.' Each trait or 'character' may be inherited in more or less independence of other traits or 'characters.'

THE ANALYSIS OF MENTAL INHERITANCE

It has been noted that the fundamental fact in mental inheritance is the relation between the germ making the parent and the germ making the child, not between the parent and child themselves. It has also been noted that inheritance is not merely of total natures vaguely, but of particular details, traits or 'characters' each more or less independently of the rest. So far as mental traits are thus inherited, we have to think of each amount of each trait as possessing some determiner in the germs and of the determiners of any one trait in different generations of germs as standing in some relation.

What is called Mendelism or Mendelian inheritance (after its discoverer, Gregor Mendel) offers an account of certain features of these determiners and of the relations in which determiners of the same trait in successive generations of a family stand.

In its clearest and most unlimited form* this account would state that:—

First.—The determiner perpetuates itself with little or no variation. No such differences exist between the determiners of trait A in germ generation 1 and the determiners of trait A in germ generation 2, as are found between parents and children.

Second.—One fertilized ovum has either (1) no determiner for trait A (when neither germ nor ovum had one), or (2) a single determiner (when either the germ alone or the ovum alone had one) or a double determiner (when both germ and ovum had one).

Third.—The germ cells later developing from one fertilized ovum will in these three cases respectively be (1) all without the determiner, or (2) half with it and half without it, or (3) all with it.

Fourth.—Any difference between one man and another in original nature is reducible to the presence or absence of such determiners. Variation in them or blends of one with another are not required.

*There are other possible views of the implications of the Mendelian results with respect to the nature of the determiners, their behavior and their relations to the 'characters' that result from them. There are also other views of the universality of the laws given here.

Fifth.—A 'character' or trait or feature of an individual may be a unit character or a multiple character. A unit character is one which is caused by the presence of one single or double determiner in the fertilized ovum from which the individual developed (positive unit character), or by the absence of one single or double determiner (negative unit character). A multiple character is one caused (1) by the presence of determiners of more than one sort, or (2) by the absence of determiners of more than one sort, or (3) by the presence of one or more and the absence of one or more of a different sort.

These laws may be illustrated as follows: Call the condition of a character, say A, in the individual which is produced by a single determiner in the fertilized ovum from which he springs, A *simplex*. Call the condition of the corresponding character in the individual which is produced by a double determiner in the fertilized ovum, A *duplex*.

Suppose co-operativeness to be a positive unit character. Call it C, and call its determiner c. Then the condition in the parents may be:—(1) both with C duplex, (2) one with C duplex and one with C simplex, (3) both with C simplex, (4) one with C duplex and one without C, (5) one with C simplex and one without C, (6) both without C. The results will be as follows:

(1) If all ova have c and all sperms have c	the unions will have 2c	and the offspring will have C duplex.
(2) If all ova have c and half of the sperms have c, or if half of the ova have c and all sperms have c	half of the unions will have 2c	and the offspring will have C duplex.
	half of the unions will have 1c	and the offspring will have C simplex.
(3) If half of the ova have c and half of the sperms have c	one-fourth of the unions will have 2c	and the offspring will have C duplex.
	one-half of the unions will have 1c	and the offspring will have C simplex.
	one-fourth of the unions will have no c	and the offspring will have no C.

(4) If all ova have c and no sperms have c or if all sperms have c and no ova have c	the unions will have 1c	and the offspring will have C simplex.
(5) If half of the ova have c and no sperms have c or if none of the ova have c and half of the sperms have c	half of the unions will have 1c	and the offspring will have C simplex.
	half of the unions will have no c	and the offspring will have no C.
(6) If no ova have c and no sperms have c	none of the unions will have c	and none of the offspring will have C.

These results seem at first sight in sharp contrast with the so common fact of blended inheritance. Blended inheritance denotes the appearance, in offspring from parents possessing A and B respectively, of conditions tending toward $\frac{A+B}{2}$, these conditions varying amongst the offspring around one central tendency and not falling at all into two or three distinct groups. But the gross facts of blended inheritance in a trait could perfectly well come by the action of the strict Mendelian laws if the trait depended on a sufficiently large number of different determiners.

It is interesting to see the resulting gradations if two, three, four, and more determiners produce each an amount of the same mental trait. Suppose, for example, that a, b, and c are determiners of intellect, aa, bb, and cc being the normal double determiners. Suppose aa to produce in the individual an amount of intellect represented by 2q, bb to produce an amount represented by 4q, and cc to produce an amount represented by 8q. Suppose a, b, and c to produce each approximately half as much as aa, bb, and cc, respectively.

The union of a germ and ovum could have:—

	o	giving	o	degrees	of	intellect	in	the	individual,
or	a	"	1q	"	"	"	"	"	
"	b	"	2q	"	"	"	"	"	
"	c	"	4q	"	"	"	"	"	
"	aa	"	2q	"	"	"	"	"	
"	ab	"	3q	"	"	"	"	"	
"	ac	"	5q	"	"	"	"	"	
"	bb	"	4q	"	"	"	"	"	
"	bc	"	6q	"	"	"	"	"	
"	cc	"	8q	"	"	"	"	"	
"	aab	"	4q	"	"	"	"	"	
"	aac	"	6q	"	"	"	"	"	
"	abb	"	5q	"	"	"	"	"	

or any one of many other combinations of the determiners, such as aabbcc, aabbc, abbc, etc. The resulting degrees of intellect would vary from 0 to 14q by steps of one.

The offspring of two individuals would then commonly show with respect to intellect as a whole a great variety of degrees as in blended inheritance, though the result with respect to the effect of any one determiner followed pure Mendelian principles. For example, suppose one individual to possess the character due to a, duplex; that due to c, simplex; and to lack that due to b. His intellect will be 6. Suppose another to possess the character due to a, simplex; that due to c, duplex; and that due to b, simplex. His intellect will be 11. On Mendelian principles the germs of the first individual will comprise germs with a and c and germs with a alone. The germs of the second individual will comprise germs with a, c, and b, germs with a and c, and germs with b and c. The unions will comprise ac + acb, ac + ac, ac + bc, a + acb, a + ac, and a + bc. The resulting intellects will comprise degrees 12, 10, 11, 8, 6 and 7. Mendelian principles applied to the elementary factors of heredity might then give, with respect to the total traits compounded of these factors, a continuous gradation and all the complexities of mental inheritance that observation finds.

Many features of animals and plants have been more or less perfectly analyzed into their unit characters, and the determiners of these have been got under control in breeding, so that strains with a desirable determiner present in all the germs have been established. Undesired determiners have been weeded out of all the germs of certain individuals, so that the offspring from any two of these individuals are sure to be free of the undesired character.

The Mendelian ideal is to represent all of each man's original nature by a list of unit characters, to refer each of these to its determiner, and to determine the future of the race by arranging selection and elimination of the determiners. This ideal is, in almost all ways, a great advance over the older plan of representing each man by a list of qualities chosen largely by historical accidents, referring these vaguely to the germ plasm, and seeking to determine the race's future by selection and elimination of individuals.

In one respect, however, it may not be better. This concerns its insistence on the invariability of the determiners from generation to generation, and within each fraternity of germs. If this invariability is perfect, any unit character can appear in only two forms or degrees, one corresponding to the presence of a double determiner in the united sperm and ovum whence the individual developed, and one corresponding to the presence of a single determiner. But if this requirement is made strictly, it is hard to find any unit characters. Curliness of hair and brownness of eye, for instance, seem to be far from constant. The former may be separated fairly clearly from straightness,* but within its own range it varies from waviness to extreme kinkiness.

As a result of the difficulty of finding any traits appearing in only two degrees each substantially invariable, many students

*Though even this may be doubted.

of heredity would frankly admit that the determiners did vary somewhat and that one determiner could blend with another similar determiner.

Mendelian Inheritance in the Case of Intellectual and Moral Traits

The possibility of analyzing mental traits into unit characters and controlling their appearance in the human species by breeding so as to get the determiners of desired traits established in all the germs, and to get all the germs freed from the determiners of undesired ones, is very attractive. If bad temper in man is as simple a compound of unit characters as hornedness in cattle or color in mice, we may hope to raise a race of assuredly good-tempered men!

It has been suggested [Davenport '10, p. 14f.] that imbecility is a unit character, depending on the absence of a certain determiner. "That imbecility is due to the absence of some definite simple factor is indicated by the simplicity of its method of inheritance. Two imbecile parents, whether related or not, have only imbecile offspring. Barr gives us such data as the following from his experience. A feeble-minded man of 38 has a delicate wife who in 20 years has borne him 19 defective children. A feeble-minded epileptic mother and an irresponsible father have 7 idiotic and imbecile children. The L. family numbers 7 persons, both parents and all 5 children imbecile. Among the "Family Records" I have been collecting there occurs the R. family where A (insane) marries in succession two mentally weak wives and has 13 children, all mentally weak. . . . . In a case described by Bennett, a defective father and an imbecile mother have 7 children all more or less mentally and morally defective. There is, so far as I am aware, no case on record where two imbecile parents have produced a normal child."

I fear, however, that the inheritance of imbecility will be found by no means so simple as Dr. Davenport hopes. If it were due to the absence of some definite simple factor, there should be some clear division of intellects into those from germs totally lacking this determiner, those from germs having a single determiner, and those from germs having the normal double determiner. There is certainly no such clear division and it is very doubtful if there is any greater division between imbecile and not quite imbecile, than between the latter and an intellect a trifle higher, and so on up to and beyond individuals of average intellect. The condition of the children in the families mentioned above should certainly be made the subject of very careful measurements before it is assumed that they are all sharply distinct from the offspring of parents whose germs possess the 'intelligence' determiner. Rough estimates are very unsafe.

It seems probable that two imbecile parents produce widely varying offspring including some more imbecile than they and some far higher than they on the intellectual scale. Richardson ['02, p. 9] writes: "Imbeciles who have been impregnated by imbeciles have produced normal children as a rule in the few cases recorded, and in my experience I know of a case where a feeble-minded boy of eighteen impregnated a feeble-minded girl of sixteen, producing a perfectly normal child."

The facts reported by Rosanoff and Orr ['11] have been interpreted as evidence that the neuropathic constitution is a unit-character due to the absence of a determiner for mental health and balance. I should like to be convinced of this, since its truth would imply the possibility of eliminating the neuropathic constitution entirely from the world by a very few generations of selective breeding. I fear, however, that neuropathy is a widely variable fact, existing in many degrees, due to the absence (or presence, or presence and absence) of many determiners, or of many intensities of one determiner.

If it were due to the absence of one determiner the offspring of parents both neuropathic should be all neuropathic and all somewhat equally so. Now the facts, as Rosanoff and Orr report them in the case of the offspring of such matings who lived beyond infancy and were traced, are as shown below.

In State hospital	16 individuals
Attack of depression, with suicidal tendency	1 individual
Fainting spells	1 "
Very peculiar, eccentric	1 "
Has been very nervous and melancholy for last two years	1 "
Had nervous break-down, hypochondriacal	1 "
Eccentric, not well-balanced	1 "
Bad tantrums, bad temper	1 "
Eccentric, bad temper	2 individuals
Quick tempered, severe headaches	1 individual
Temper easily excited, "nags repeatedly"	1 "
Nervous temperament, easily excited, easily upset	1 individual
Nervous temperament, excitable	1 "
Nervous temperament, high-strung	1 "
Nervous temperament, fretter	1 "
Very excitable, high-strung	1 "
Nervous temperament	1 "
Defective, cranial malformation	
Sick headaches	2 individuals
Very alcoholic, very bad temper	1 individual
Quick tempered, not very bright	1 "
Stutterer, nervous	1 "
Fidgety, cannot keep still	1 "
Bad temper	1 "
Alcoholic, lazy, indolent	1 "
Very stubborn	1 "
Inferior, "slow"	1 "
Very alcoholic, not very bright	1 "
Peculiar temper	1 "
Asthma	1 "
"Easy-going," weighs 300 lbs	1 "
Apparently normal, 20 years old	1 "
Normal (ages 22, 20, 18, 17, 15, 15, 8)	7 individuals

They range from clearly insane individuals, through bad nervous break-downs and marked eccentricity, to individuals about whom nothing worse than excitability, asthma, a bad temper, stubbornness, hard drinking or "easygoingness" could be found, and to those who were free from any evidence whatsoever of mental instability. They do not at all fit the assumption that human beings are sharply divided into those with and those without neuropathic tendencies. They fit perfectly the hypothesis that human beings vary in respect to neuropathic tendencies by gradual steps from the extremely insane to the extremely well-balanced. But if the latter is the case, the neuropathic tendency cannot be due to the absence of one determiner. Rosanoff and Orr, indeed, incidentally admit that it is not ("It seems necessary to assume that the normal development and function of the nervous system is dependent not upon a single unit determiner in the germ plasm, but upon a group of determiners" [p. 230]). This admission makes their previous defence of Mendelian inheritance of neuropathic constitution rather pointless.

Mental traits are certainly not as a rule unit characters or the results, each of two or three cooperating unit characters. On the contrary, most of them seem to be the results of very many unit characters. It will be shown in Chapter XIV that no case is known of a mental trait appearing in three sharply defined degrees,—0, moderate, and full amount,—with clear gaps between, as a unit character should. Almost, if not quite, all mental traits, so far as they are due to original nature, appear in many different degrees, each varying very slightly from the next. But this gradation in a trait's amount necessitates the existence of many determiners to produce it, if it is to be produced by determiners that do not themselves vary greatly in force. So, for intellectual and moral traits, the task of analysis into unit characters and attribution to invariable determiners seem likely to be very, very difficult.

It is not necessary to decide whether the facts of mental inheritance can be explained by constant determiners or require variable determiners; nor whether, if the latter be the case, two variable determiners of the same trait may blend. The question must be left for investigation by students from both sides. The chief present value of the Mendelian facts and hypotheses to students of intellectual and moral traits is as an encouragement to the more exact description and analysis of the original natures of individual men, and to more exact measurements of individual resemblances.

CHAPTER XII

The Influence of Maturity

No competent student doubts that in certain mental traits maturity or inner mental growth causes one individual to differ year by year from his former self, irrespective of all training. The same force necessarily accounts for some of the differences found between children of different degrees of mental maturity. If by a miracle a hundred children could be found who were alike in sex, ancestry and training, but who were divided into two groups by a difference in the extent to which the original impetus to mental development had run its course, the groups would differ, in at least certain traits, in accordance with this difference in stage of growth or maturity.

About the magnitude of the influence of maturity there is, however, a wide range of opinion, from that which would expect children in the same stage of growth to be all closely alike and all very different from children in a later stage of growth, regardless of differences in their ancestry and training, to that which would expect children of the same ancestry and training to be all very much alike, regardless of differences in stage of growth.

The study of the facts is made difficult by the absence of any exact measure of maturity, that is, of the extent to which the original impetus to mental development has run its course. Length of life is the measure which has been used, but chronological age is not identical with physiological maturity and neither of these two is identical with mental maturity. An individual's degree of mental maturity cannot be inferred from his age. In the long run, however the central tendencies of children of the ages 6, 7, 8, etc., will represent the central tendencies of successive stages in mental growth. And in any case such comparisons of children at different ages are the only measurements available in support of conclusions about the influence of maturity.

CHANGES IN MENTAL TRAITS WITH AGE

Some of the best known and most commended studies in educational psychology deal with the differences in mental traits between children of different ages. The most extensive and also the most painstaking study of this sort is Dr. Gilbert's *Researches on the Physical and Mental Development of School Children* ['94]. A fairly careful examination of its method and results will be our best introduction to the general problems of the chapter.

Dr. Gilbert made a number of measurements of both physical and mental traits in boys and girls from six to seventeen years old. The mental traits were:

1. Delicacy of discrimination of weight ('Muscle-Sense'). (Ten weights, identical in shape and size, but weighing 84 grams, 86 grams, etc., were set before a child and he was asked to sort out all those which seemed to him to be of exactly the same weight as the 82-gram one (which was marked by a white dot). Delicacy of discrimination was then measured inversely by the difference in weight of the weights thought to be identical.)

2. Delicacy of discrimination of color. (A series of reds varying progressively in darkness were used as the weights were in 1.)

3. Force of suggestion. (Measured by the amount a child overestimated a weight small in size compared with the same weight made much larger.)

4. Voluntary motor ability. (The number of taps made with the finger in 45 seconds.)

5. Fatigue. (Let T = the number of taps made in the

first 5 seconds of a trial for 45 seconds. Let L = the number of taps made in the last 5 seconds. Let $F = \frac{T-L}{T}$. F was the measure of fatigue used.)

6. Reaction time. (Measured by the time taken to see a signal and react by pressing down a key.)

7. Reaction with discrimination and choice. (Measured by the time taken to see that the signal was blue and not red and to react by pressing down a key.)

The essential results of Gilbert's study are given in Table 35. It tells with fair accuracy the median ability of every such group as 'girls from 9 years 0 months to 9 years 11 months inclusive,' and the variability of every such group in four of the traits measured. In three traits the variability is given only for boys and girls together.

TABLE 35.

THE CENTRAL TENDENCIES AND VARIABILITIES OF CHILDREN OF DIFFERENT AGES IN DISCRIMINATION OF WEIGHT, DISCRIMINATION OF SHADES OF RED, RESISTANCE TO THE SIZE-WEIGHT ILLUSION, ETC.

M.V. = in all cases the average deviation of the individual children from the median child of that year-age.

B = in all cases, boys.

G = in all cases, girls.

D wt = the number of grams difference required in order that the median child should perceive the difference.

D col. = the smallest number of differences (not objectively defined) required in order that the median child should perceive the difference. For a statement of the nature of the differences see Gilbert, '94, p. 42 f.

H = the median difference in grams between the two weights (of equal size) chosen as equal respectively to two blocks, each weighing 55 grams and being 2.8 cm. thick, but being 2.2 cm. and 8.2 cm. in diameter.

T = the number of taps made in the first 5 seconds of 45 by the median child.

$F = \frac{T-L}{T}$ (L being the number of taps made in the last 5 seconds of 45 by the median child).

Rs = the time, in thousandths of a second, between the movement of a disc and the making of a contact by a child who is instructed to press down the key as soon as he sees the disc move: the median child's time.

Rd = the time to see a blue surface and react to it as for Rs., no reaction being permissible if the surface shown was red instead of blue.

MUSCLE-SENSE

Age.	*Dwt.(B+G)*	*MV(B+G)*	*Dwt.(B)*	*Dwt.(G)*
6	14.8	5.2	13.0	16.8
7	13.6	4.4	13.2	13.2
8	11.4	4.6	12.2	11.0
9	10.0	4.4	10.2	10.0
10	8.8	4.4	8.6	9.2
11	8.6	3.8	10.2	7.6
12	7.2	3.0	7.6	7.6
13	5.4	3.0	6.0	5.6
14	5.6	3.0	5.2	7.2
15	6.8	2.2	6.2	7.2
16	6.6	2.4	6.0	6.8
17	5.8	2.6	6.0	6.4

SENSITIVENESS TO COLOR-DIFFERENCES

Age.	*Dcol.(B+G)*	*MV(B+G)*	*Dcol.(B)*	*Dcol.(G)*
6	9.6	1.8	8.3	9.6
7	9.0	2.1	8.3	9.6
8	8.3	2.3	9.6	7.0
9	6.3	2.2	6.1	6.6
10	5.4	1.9	6.0	5.2
11	5.4	1.7	6.0	4.9
12	5.1	1.5	4.8	5.1
13	4.6	1.7	5.2	4.1
14	4.7	1.4	4.8	4.6
15	4.4	1.1	4.1	4.6
16	4.3	1.3	4.3	4.0
17	3.9	1.4	4.0	4.9

FORCE OF SUGGESTION

Age.	*H(B+G)*	*MV(B+G)*	*H(B)*	*H(G)*
6	42.0	17.0	43.5	42.5
7	45.0	15.5	43.5	43.5
8	47.5	13.5	45.0	49.5
9	50.0	10.5	50.0	49.5
10	43.5	12.5	40.0	44.0
11	40.0	11.5	38.5	40.0
12	40.5	9.0	38.0	41.0
13	38.0	9.0	37.0	38.0
14	34.5	9.5	31.0	33.5
15	35.0	10.0	33.0	38.0
16	34.5	10.0	32.0	38.5
17	27.0	12.0	25.0	31.0

Voluntary Motor Ability

Age	*T(B+G)*	*MV(B+G)*	*T(B)*	*MV(B)*	*T(G)*	*MV(G)*
6	20.8	2.4	21.0	2.5	19.7	2.5
7	22.5	2.9	22.8	2.7	21.2	2.5
8	24.4	2.9	24.9	3.4	23.9	2.2
9	25.4	2.5	25.8	2.5	25.0	2.9
10	27.0	2.8	27.7	2.6	26.9	2.8
11	29.0	3.3	29.7	3.2	27.8	3.0
12	29.9	3.3	30.3	3.1	29.6	3.0
13	28.9	2.8	29.8	3.0	28.1	3.3
14	30.0	3.6	31.2	3.2	28.0	3.4
15	31.1	3.0	31.3	2.6	29.8	3.2
16	32.1	3.3	33.0	3.0	31.8	3.4
17	33.8	2.9	35.0	2.4	31.5	2.3

Fatigue

Age.	*F(B+G)*	*MV(B+G)*	*F(B)*	*MV(B)*	*F(G)*	*MV(G)*
6	21.4	8.1	22.8	9.4	21.3	7.0
7	21.0	8.9	22.5	9.7	20.2	6.7
8	24.0	7.3	24.7	8.3	23.3	7.1
9	21.0	7.1	22.5	6.7	20.7	7.8
10	22.0	7.5	22.7	7.8	19.0	7.1
11	20.0	6.2	20.3	6.5	18.0	5.5
12	16.0	6.3	18.0	6.0	14.0	6.7
13	14.5	6.4	15.8	6.7	14.7	5.8
14	14.0	6.5	17.8	6.2	12.0	6.1
15	12.7	5.8	13.8	4.9	11.5	5.7
16	14.7	5.2	15.3	4.6	11.7	5.6
17	13.8	5.3	14.5	6.3	13.5	4.3

Reaction Time

Age.	*Rs(B+G)*	*MV(B+G)*	*Rs(B)*	*MV(B)*	*Rs(G)*	*MV(G)*
6	295	50	282	46	295	54
7	292	55	267	46	315	52
8	262	39	245	39	260	31
9	250	41	243	54	255	49
10	215	36	210	26	225	43
11	195	34	185	31	206	34
12	187	31	178	27	198	35
13	187	30	178	29	205	35
14	180	29	180	30	187	30
15	172	27	167	23	189	27
16	155	23	147	16	172	26
17	155	33	147	19	163	26

REACTION WITH DISCRIMINATION AND CHOICE

Age.	Rd(B+G)	MV(B+G)	Rd(B)	MV(B)	Rd(G)	MV(G)
6	525	60	535	53	510	65
7	530	81	490	88	528	94
8	478	65	480	57	475	55
9	450	68	445	63	460	72
10	410	49	400	49	415	45
11	385	58	387	58	388	57
12	370	55	385	60	370	49
13	395	58	360	51	415	55
14	365	49	367	45	355	54
15	335	49	311	55	345	38
16	325	43	315	39	350	39
17	312	40	305	35	315	44

Just what do these median abilities of Table 35 mean? Just what do the differences between those for six and seven, seven and eight, etc., tell us about the development of mental traits in life? Just what do we learn about human nature from these comparisons of the capacities of children of different ages?

It is clear that an alteration in any mental trait in any individual with age might be due to the mere maturing of some characteristic of original nature or might be the creation of some environmental force. The educational inferences would be exactly opposite in the two cases. In the former we should say: This change comes as a gift from nature which we may not be able to refuse without damaging general growth. It is given as the partial basis and starting point for education. We do not have to try to get it. In the latter case we should say: This change comes as the earnings of training. It is a product of education. With a different training it might be absent. We may lack or possess it as we choose.

Moreover, in the case of many measurements of mental traits, for instance those quoted, the change due to an individual's age would be possibly due not only to the maturing of the trait or the influence of training upon it, but also to the influence of both maturity and training upon the ability to understand and the wish to follow instructions and the ambition to do well in tests. This complex of traits we may call general ability in tests. It is even conceivable that the last factor was the sole cause of all the changes quoted above.

As a matter of fact all three of these factors are involved in most of the changes of mental traits with age. Even if the changes are due directly to outside forces, in the form of the experiences of life and training, maturity may still count as a force co-operating with these or furnishing the conditions in the individual which permit their action on him to produce the mental changes in question. On the other hand, mere inner growth, no matter how potent, requires usually some stimuli from without. A child grows mentally in some kind of a world of experience, forming some habits. Only in thought can the contribution of his inner impulsions be separated off from the contribution of the outside stimuli by which the inner impulsions are roused to action. Furthermore, a mental test with children almost always measures somewhat general powers of comprehension as well as the special power of sensation, memory or the like that is its ostensible object.

Hence mere knowledge of age differences in a mental trait, without knowledge of how their causation is distributed amongst these three factors, does not allow us to estimate the amount of influence of maturity in determining an individual's status in respect to the trait. The total age change must at least be divided between mere maturity and the added training which has accompanied maturity. Until this is done we cannot progress far beyond the vague commonplaces of common observation.

THE DIFFICULTIES IN INFERRING CHANGES IN INDIVIDUALS WITH AGE FROM DIFFERENCES BETWEEN OLD AND YOUNG INDIVIDUALS

So far upon the supposition that by changes in mental traits with age, we mean changes in the same individuals measured at different ages. The average change would then be the average of the changes in all the individuals studied. But in the studies that have been reported, the difference between the figures for, say, ten and eleven years, is not the average of the changes of all the individuals studied and need not in any real way describe them.

For (1) the difference between the average of a group at ten and of the same group at eleven years does not describe the real individual changes; and (2) when we measure ten-and eleven-year-olds as we find them in school or elsewhere, we can not be sure that the eleven-year-olds represent what the ten-year-olds will become.

The first point will be made clear by the following illustration. Suppose that eighteen boys showed at the age of ten and a half years the abilities in some mental trait denoted by the measures in the first column and made the gains during the next year shown by the figures in the second column, their consequent records at eleven and a half years being given in the third column. (Case 1.)

Case 1			Case 2		
Ability at 10½	Change	Ability at 11½	Ability at 10½	Change	Ability at 11½
2	5	7	2	0	2
2	5	7	2	0	2
3	4	7	3	1	4
4	3	7	4	0	4
4	4	8	4	1	5
5	4	9	5	3	8
5	1	6	5	1	6
6	3	9	6	1	7
6	3	9	6	1	7
6	1	7	6	3	9
6	1	7	6	3	9
7	1	8	7	1	8
7	3	10	7	4	11
7	1	8	7	4	11
8	0	8	8	3	11
9	1	10	9	4	13
9	0	9	9	5	14
11	0	11	11	5	16
Avg. 5.94	2.22	8.16	Avg. 5.94	2.22	8.16

If instead of this complete record we had simply the figures: 10½ years, Av. 5.94; 11½ years, Av. 8.16; Change in average ability, 2.22, we should lack the essential features of our fact; viz., (1) the variability of the changes and (2) the antagonism between ability at ten and a half years and growth during the next year. There is an almost inevitable tendency, when a single figure is given to represent change, to fancy that all children show exactly or nearly that amount of change. This is of course never true. Rate of change as well as absolute ability is variable. And it is precisely in relating the different degrees of progress found in individuals to their original capacities and individual circumstances, that educational insight will accrue. The real individual changes may often prove to be a partial function of the amount of ability already acquired, as in our illustration. The mere change in average ability given above could have come as well from a condition, shown in Case 2, just opposite in this respect to that of Case 1.

In Case 2, the better a boy is at ten and a half years the more he gains; whereas in Case 1, the better he was the less he gained. Case 2 would, I venture to prophesy, be the fact in the progress with age of real mental efficiency, while with physical growth from thirteen to eighteen we should have something like Case 1, the children who had matured early and so attained high stations in stature growing little, while those who matured slowly would keep on growing at a fair rate. In brief, the growth of averages does not accurately describe, and may positively misrepresent, the real growth of the individuals in the group.

Our second point was that the eleven-year-olds tested need not represent what the ten-year-olds would become. The average changes stated in the quotations at the beginning of this chapter were obtained from facts like the following: Ten-year-olds *A, B, C, D, E, F, G, H,* etc., give an average x; eleven-year-olds, *L, M, N, O, P,* etc., give an average y. The change in average ability is $y - x$. The individuals of the two groups not being identical, the chance is given for the fallacy of selection to run riot. The eleven-twelve-year-olds certainly represent only those ten-eleven-year-olds who will live; in any test given in schools they represent only the ten-eleven-year-olds who will continue in that type of school. Now if one measures a mental trait in elementary school children he gets for different ages something like the following figures:—12-year-olds, 100; 13-year-olds, 90; 14-year-olds, 70; 15-year-olds, 30.

Nobody can imagine that the fifteen-year-olds here would give anything like a fair sampling of what the twelve-year-olds would become. The brightest twelve-year-olds pass out of the grammar school before they are fifteen. Some mental defectives leave for special institutions. Some moral defectives leave for reform schools or the free life of thievery and trampdom. Some children leave school to go to work. If we fill up our quota of fifteen-year-olds by adding 70 from high school pupils we jump from the frying pan into the fire, for these are a selection of the brighter, the more ambitious, and the more intellectually inclined.

I conclude, therefore, that the development of mental traits with age has not been and can not be adequately measured by such studies as those quoted. To measure it we must repeat measurements upon the same individuals and for all purposes of inference preserve intact each of the individual changes. In connection with each of them account must be taken of the training which the individual in question has undergone.

What measurements we do have may serve, however, to correct two errors of common opinion. The notion that the increases in ability due to a given amount of progress toward maturity are closely alike for all children save the so-called "abnormally precocious" or "retarded" is false. The same fraction of the total inner development, from zero to adult ability, will produce very unequal results in different children. Inner growth acts differentially according to the original nature that is growing.

The notion that maturity is the main factor in the differences found amongst school children, so that grading and methods of teaching should be fitted closely to 'stage of growth,' is also false. It is by no means very hard to find seven-year-olds who can do intellectual work at which one in twenty seventeen-year-olds would fail. Although the influence of inner growth in causing individual differences cannot be measured from the data at hand, an *upper limit* for it can be set. Take discrimination of weight as a sample case. Since early age differences are in part due to training and since training acts here in the same direction as does maturity, the average inner growth from, say, ten to seventeen must produce *less* than the average difference found between ten-year-olds and seventeen-year-olds. Since, in Gilbert's study, the seventeen-year-olds and ten-year-olds both come from school pupils, including pupils in the high school, the seventeen-year-olds represent at least as high ranking pupils in mental respects as the ten-year-olds would become. So the effect of average inner growth from ten to seventeen is at the outside a reduction in the least noticeable difference from 8.8 to 5.8 grams. But many ten-year-olds noticed a difference of 5.8 grams or less. Gilbert's 4.4 (for the average deviation of the individual ten-year-olds from their C. T.) would put 30 per cent of them above the average seventeen-year-old unless the distribution of individual differences among the ten-year-olds is markedly eccentric. And within the ten-year-olds there is a range of variation at least five times as great as that between the average ten-year-old and the average seventeen-year-old. The range cannot be less than four times 4.4 g. unless the distribution (surface of frequency) is of a form never found in measurements of discrimination of weight. Hence even the topmost limit for the average effect of these seven years' maturity is surely less than one sixth of the effect of the extreme differences in ancestry and training upon children of the same age.

CHAPTER XIII

The Influence of the Environment

DIFFICULTIES IN ESTIMATING THE AMOUNT OF INFLUENCE OF THE ENVIRONMENT

The questions suggested by the title of this chapter include the effects on individuals of every environmental force, including all the agencies for intellectual and moral education. Precise quantitative answers can be given to hardly any of them.

Theoretically, there is no impossibility. Once we have estimated the original nature of a man or group of men, we have simply to note the mental changes consequent upon this or that change in climate, food, school training, friendship, sermon, occupation, etc. Practically, the complexity of the action of physical and human influences upon intellect and character hampers scientific study and favors guesswork. The environment includes a practical infinitude of different causes; these act differently upon different types of original nature and at different ages and with different co-operating circumstances; in many cases their action is very complex and must be observed over long intervals of time. Indeed it has been common to deny even the possibility of a science of the dynamics of human nature and to remain content with the haphazard opinions of novelists, proverb makers and village wise men.

Moreover, it is only by the utmost ingenuity and watchfulness that studies of changes in human nature can be freed from a characteristic fallacy—that of attributing to training facts which are really due to original nature or to selection. For instance, college graduates are found to have a much greater likelihood of being elected to Congress than other men have. Therefore it is said that a college education causes to some extent political success. But it is clear that even before they went to college the group of youth who did go were different from those who did not. Their later election to Congress may as well have been due to the mental traits which they possessed by birth or otherwise and which caused their inclusion in the class 'boys who go to college' as to any changes produced in them by the college training itself. In other words, that they were the class *selected* by the college is as important a fact as that they were the class trained by it.

Again it is said: "Who can doubt the enormous disciplinary value of the study of Latin and Greek when we see the admirable intellects of the men so trained in the English universities?" But being born from the class whose children go to the university of itself ensures to an individual uncommon mental ability.

To avoid this confusion of causes which train with those which select is extremely hard. Any class of individuals studied because they have been subjected to a certain training is almost sure to be a class not only trained by but also selected by that training. Suppose that one wishes to study the influence of a high-school course, or that of the classical as opposed to the scientific course, or that of training in independent research, or that of immoral surroundings. High school graduates are but one-fifth of grammar school graduates; and no one would claim that they represent an entirely random picking therefrom. They are surely selected for better birth, better abilities and better ideals. Again, in most high schools the graduate of the classical course represents not only a different training, but also a different selection, commonly a superior selection.* So also scientific men are a class resulting not only from the training given by research work, but also from the selection of those eager to do and fitted to do that

*This apparently is becoming less common every year.

work. Children brought up in a morally bad environment are almost sure to be of morally inferior ancestry. The ordinary arrangement of social and educational careers rarely presents us with convenient cases of similar natures, some with, some without, the form of training under consideration.

The difficulty of eliminating the influence of selection is no excuse for its neglect. Yet one may hunt through thousands of pages of discussions of the influence of certain studies, school systems, schemes of culture, religious beliefs, etc., without finding a hint of its recognition.

Either because of the general complexity of environmental influences upon any mental trait and the mixture of selective with formative influences or because of the infrequency of scientific habits and ideals in students of sociology and education, there are few facts of sufficient security and precision to be quoted. Only rarely has educational science progressed beyond the reasoned opinions of more or less capable judges. We have our beliefs about the causal relations between a hot climate and indolence, necessity and invention, lack of parental control and crime, religious training and morality, etc., but we can not be said to know these influences with adequate surety or to have any knowledge whatever of their precise amount.

A refusal to believe insecure opinions about the influence of differences in training in producing differences in human individuals does not at all imply disbelief in their influence. Such would be absurd. When the original natures are the same, *every* difference that the individuals later show must be due to differences in the outside forces operating upon them. And any difference in outside forces always has its effect. No man is left unchanged by even the very least of the environmental forces that act upon him. Men are the creatures of circumstance. But they are creations whose final patterns are determined in part by sex, race, ancestry and conditions of origin. Circumstances alter natures, but the alterations vary with the nature altered. It is precisely because common opinions have thought verbally in terms of '*man-training-product of training*,' instead of concretely in terms of *'men-training-products,—each of an individual's nature in interaction with his training,'* that a sound science of the influence of the environment has hardly been begun.

One of the best services such a science can render is to guard its students against such verbal plausibilities. For example, knowledge is not proportional to opportunity in the sense that an individual's degree of knowledge can be foretold from his degree of opportunity. Wealth does not create wealth in the sense that what a man will have can be estimated from what he now has. A good home does not make good children in the sense of doing so always and in proportion to its goodness. Being treated like slaves may not debase all and never debases all alike. The product of the environment is always a result of two variables, it and the man's nature.

Two of the corollaries of this axiom are of special significance. The first is that the environmental stimulus adequate to arouse a certain power or ideal or habit in one man may be hopelessly inadequate to do so in another. Washing bottles in a drug-shop was, if a common story is true, adequate to decide Faraday's career, and the voyage on the Beagle is reputed to have made Darwin a naturalist for life. But if all the youth of the land were put to work in drug-shops and later sent on scientific expeditions, the result would not be a million Faradays and Darwins, or even a million chemists and naturalists. All that one man may need to be free is a vote; but even a long education in self-direction may be inadequate for another. Being told a few words suffices to secure the habit of reading in one child, while the child beside him remains illiterate after two years of careful tuition. The amount of stimulus required in some cases is so infinitesimal that the power seems to spring absolutely from the man himself. In other men no agency is found potent enough to arouse a trace of the desired result.

The second corollary is that each man in part selects his own environment. The boy turns his eyes from the book. Even if his eyes attend to it, his mind does not. Even if for the time he lets it move him, it may be disregarded in memory. That connection which brings satisfaction to one man and is thereby given power over him, may disgust another nature and so be repudiated by it. As this world's nature selects for survival those animals which are adapted to live in it, so any individual selects, by action, attention, memory and satisfaction, the features of the environment which are to survive as determinants of his intellect and character.

Common opinion and the older literature of sociology and education neglected the differential action of the environment in accord with the nature it acted on, but it would be possible for a student, enamored of the simplicity of the explanation of all men's differences by differences in their original make-up, to neglect equally obvious facts of another sort. He might be tempted to claim that, since the features of civilization,—the acts, words, books, customs, and institutions of men,—have been invented and perpetuated by human natures, and since consequently the environment in all important respects is itself due to original nature,—therefore original nature is at bottom the cause of almost all of human destinies. "A people gets as good government as it deserves; a race has the environment its own nature has found and chosen: a man in essential matters is treated as his nature decides." So he might carelessly claim.

Many important features of the environment are thus due to the original nature of the human race as a whole, but no one man's nature and, under modern conditions, no one nation's or race's is similarly responsible for the particular environment that it meets. Forces set in motion by others play upon it. At the best it can select only negatively by disregard, and at the worst it may be molded directly against nature.

Even when it is known, and with some precision, that a given difference is due to some difference in training, there may be doubt or total ignorance as to what difference in training caused it. And even when it is known that a given difference in training has been operative and has produced an effect, there may be doubt or ignorance about what the effect is.

Illustrations of the former case are abundant in history. History is in fact largely a record of unexplained changes in human nature. Nearly all the intellectual and moral differences between the modern English, French, or Germans and their barbarous ancestors of two thousand years ago are due to differences in environment. The original natures of the stocks may have altered somewhat during that time, but surely not much. Our thoughts and ways of thinking and our habits, customs and ideals have been and are being made very unlike those of our ancestors by some outside forces. But what the forces were and how each contributed to the result is not known.

Illustrations of the latter case form a large proportion of the facts studied under the vague rubric of education. Such and such children have gone to school, they have been taught by such and such teachers, using this and that method, at a cost of so many dollars, with aid of a material plant worth so much; but what has come of it all, no cautious thinker would dare say. What has been and is being done to children in schools is more or less well described in official and private records, but what happens in children as its consequence is largely unknown.

So much for the attitude in which a student of human nature must approach the problems of the effect of different environments on identical natures, of the effect of the same environment on different natures, and of the effect of the endless different co-operations of environments and natures.

MEASUREMENTS OF THE INFLUENCE OF THE ENVIRONMENT

I shall report four samples of studies of the influence of the environment upon intellect and character. The first is Galton's *History of Twins* ['83], a study of the amount of its influence in comparison with that of original nature. The second, from Cattell's *Statistical Study of American Men of Science* ['06], is a study of the effect of early opportunity upon scientific achievement. The third is Rice's study of the effect of different school environments upon ability in spelling and arithmetic ['97 and '02]. The fourth is a study of the effect of changing environment upon the choice of a profession by scholarly youth.

Galton collected reports from parents concerning twins who were closely similar in infancy but whose environments differed, and twins who were in infancy notably unlike, but whose environments were in all important features identical. The increase of differences in the former case and of resemblances in the latter gives a measure of the influence of the environment. The persistence of similarities in the former case and of differences in the latter gives a measure of the influence of original nature.

This evidence in the first case consists of illustrations of identical mental habits, tastes, associations of ideas and susceptibilities to mental diseases. The cases of unlikeness seem to him to be due to such alterations in the amount of energy as could be caused by illness or lowered nutrition rather than to fundamental qualities of mind.

The evidence in the case of the twenty pairs in the second group shows no exceptions to the rule that no weakening of inborn differences by similarities of nurture is observable. The following are representative parental observations:—

1. One parent says:—"They have had *exactly the same nurture* from their birth up to the present time; they are both perfectly healthy and strong, yet they are otherwise as dissimilar as two boys could be, physically, mentally, and in their emotional nature."

2. "I can answer most decidedly that the twins have been perfectly dissimilar in character, habits, and likeness from the moment of their birth to the present time, though they were nursed by the same woman, went to school together, and were never separated till the age of fifteen."

3. "They have never been separated, never the least differently treated in food, clothing, or education; both teethed at the same time, both had measles, whooping-cough, and scarlatina at the same time, and neither had any other serious illness. Both are and have been exceedingly healthy and have good

abilities, yet they differ as much from each other in mental cast as any of my family differ from another."

4. "Very dissimilar in body and mind; the one is quite retiring and slow but sure; good-tempered, but disposed to be sulky when provoked;—the other is quick, vivacious, forward, acquiring easily and forgetting soon; quick-tempered and choleric, but quickly forgetting and forgiving. They have been educated together and never separated."

5. "They were never alike either in body or mind and their dissimilarity increases daily. The external influences have been identical; they have never been separated."

6. "The two sisters are very different in ability and disposition. The one is retiring but firm and determined; she has no taste for music or drawing. The other is of an active, excitable temperament; she displays an unusual amount of quickness and talent, and is passionately fond of music and drawing. From infancy, they have been rarely separated even at school, and as children visiting their friends, they always went together.

7. "They have been treated exactly alike; both were brought up by hand; they have been under the same nurse and governess from their birth, and they are very fond of each other. Their increasing dissimilarity must be ascribed to a natural difference of mind and character, as there has been nothing in their treatment to account for it."

8. "They are as different as possible. [A minute and unsparing analysis of the characters of the two twins is given by their father, most instructive to read, but impossible to publish without the certainty of wounding the feelings of one of the twins, if these pages should chance to fall under his eyes.] They were brought up entirely by hand, that is, on cow's milk, and treated by one nurse in precisely the same manner."

9. "The home-training and influence were precisely the same, and therefore I consider the dissimilarity to be accounted for almost entirely by innate disposition and by causes over which we have no control."

10. "This case is, I should think, somewhat remarkable for dissimilarity in physique as well as for strong contrast in character. They have been unlike in body and mind throughout their lives. Both were reared in a country house, and both

were at the same schools till *aet.* 16." [Galton, '83, p. 170 f. of the reprint in Everyman's Library.]

The two lines of evidence taken together justify, in Galton's opinion, the following general statements:

"We may, therefore, broadly conclude that the only circumstance, within the range of those by which persons of similar conditions of life are affected, that is capable of producing a marked effect on the character of adults, is illness or some accident that causes physical infirmity. . . . The impression that all this leaves on the mind is one of some wonder whether nurture can do anything at all, beyond giving instruction and professional training. There is no escape from the conclusion that nature prevails enormously over nurture when the differences of nurture do not exceed what is commonly to be found among persons of the same rank of society and in the same country." [*ibid.* pp. 168 and 172.]

Even in the hands of a master, the collection of data through correspondence is inferior to direct observation and measurement. Galton was misled to believe that twins fall naturally into two groups, those much alike, and those little alike, in infancy. They do not. His correspondents may have made careless reports in other respects also. However, it will be remembered that with respect to fifty pairs of twins, objectively measured, the facts showed that the existing differences in home training had very slight effects upon the six mental abilities tested.

The conditions of nurture of men of great achievement have been studied by De Candolle ['73], Galton ['74], Jacoby ['81], Odin ['95], Ellis ['04], Cattell ['06], and others. Within any one nation such men are more likely than chance would allow to be brought up in thickly settled districts, in particular in cities, still more particularly, in the case of men of science or letters, in cities containing universities; by parents in comfortable or more than comfortable financial circumstances; and to have received a good education.

Such facts are used in Odin, and by Lester F. Ward ['06] following him, as evidence that the number of men of great achievement could be increased many times over if all men had in youth the stimulus of an intellectually active city, freedom from productive labor, and good school training. But obviously the features of the successful man's surroundings, enumerated above, might all be the secondary results of superior parentage. If men of high capacity go to live in cities, their sons will be born and reared in cities; if men of scientific and literary gifts are attracted to university cities, such cities will, on grounds of heredity alone, produce future scientific and literary men. If men of high achievement are born of men of over-average achievement, they will not be brought up by day-laborers and without education as often as chance would dictate. The parent's achievement leads forward to these environmental conditions as truly as the son's achievement leads back to them. So nothing is proved by them. From Table 36 [Cattell, '06], one can, according to his point of view, get evidence that educational opportunity counts enormously or that it counts slightly in the relative production of scientific men by the different states of this country.

For instance, the most striking fact is the high position of New England and the low position of the southern states. The advocate of great influence from opportunity may plausibly say that from 1870 to 1885, the time when most of the men in question were from 10 to 30 years of age, life in New England meant cities, schools, books, lectures, the personal example of scholarly men and freedom from poverty, whereas life in the south meant the reverse. But the advocate of little influence may retort that Maine, New Hampshire and Vermont should not, on these grounds, have been so superior to Rhode Island; or Louisiana so inferior to Texas; or Missouri so inferior to Wisconsin. He may insist that the chief gift of Massachusetts parents to their sons was the original nature which in earlier days developed commerce and manufactures, established the schools and libraries, and raised life above a daily struggle for physical necessities, while the southern planters and their servants were content to leave nature unimproved.

TABLE 36.

DISTRIBUTION OF 867 MEN OF SCIENCE BORN IN THE UNITED STATES.

	Birthplace.			Per. Million
North Alantic Division.	I.-V.*	VI.-X.*	Total.	1860.
Maine	19	10	29	46.1
New Hampshire	7	8	15	46.0
Vermont	9	9	18	57.1
Massachusetts	60	74	134	108.8
Rhode Island	4	1	5	28.6
Connecticut	26	14	40	86.9
New York	99	84	183	47.2
New Jersey	9	19	28	41.6
Pennsylvania	32	34	66	22.7
South Atlantic Division.				
Delaware	0	2	2	17.8
Maryland	12	14	26	37.8
District of Columbia	1	2	3	39.9
Virginia	5	8	13	
West Virginia	1	0	1	8.8
North Carolina	1	4	5	5.0
South Carolina	2	3	5	7.1
Georgia	1	2	3	2.8
South Central Division.				
Kentucky	6	2	8	6.9
Tennessee	5	1	6	5.4
Alabama	1	1	2	2.1
Mississippi	1	0	1	1.3
Louisiana	1	0	1	1.4
Texas	0	3	3	4.9
North Central Division.				
Ohio	42	33	75	32.1
Indiana	17	11	28	20.7
Illinois	24	18	42	24.5
Michigan	12	15	27	36.0
Wisconsin	11	24	35	45.1
Minnesota	1	3	4	23.2
Iowa	6	14	20	29.6
Missouri	4	10	14	11.8
North Dakota				
South Dakota				
Nebraska	1	1	2	69.3
Kansas	5	2	7	65.3
Western Division.				
Montana				
Wyoming				
Colorado	0	3	3	87.2
New Mexico				
Arizona				
Washington	1	0	1	86.2
California	5	6	11	28.9
Alaska				
Hawaii	1	0	1	
Philippine Islands				
Total	432	435	867	27.6

*Column I.-V. gives the facts for men of the highest ability; column VI.-X. gives the facts for the 500 men of less ability.

Such a conflict of opinion is found between Professor Cattell, who gathered the data of Table 36, and Professor Woods. Some of the comments of the former are:

"The inequality in the production of scientific men in different parts of the country seems to be a forcible argument against the view of Dr. Galton and Professor Pearson that scientific performance is almost exclusively due to heredity. It is unlikely that there are such differences in family stocks as would lead one part of the country to produce a hundred times as many scientific men as other parts. The negroes may have a racial disqualification, but even this is not proved. The main factors in producing scientific and other forms of intellectual performance seem to be density of population, wealth, opportunity, institutions and social traditions and ideals. All these may be ultimately due to race, but, given the existing race, the scientific productivity of the nation can be increased in quantity, though not in quality, almost to the extent that we wish to increase it. . . . .

"My general impression is that certain aptitudes, as for mathematics and music, are mainly innate, and that kinds of character and degrees of ability are mainly innate, but that the direction of performance is mainly due to circumstances, and that the environment imposes a veto on any performance not congenial to it." ['06, pp. 734-735]

The cities in which five or more of the thousand men of science were born, are given in Table 37. [Cattell, '06, p. 738] The author of the study is cautious in estimating the beneficial results of city life. He says simply:

"Of the 866 men native to the United States, 224 were born in the cities which in 1900 had a population of more than 25,000. These places had in 1860 a population of about 4,500,000 as compared with a rural population of about 27,000,000. The urban population was about one-sixth of the rural population and produced more than a quarter of the scientific men. The urban birth rate was 50 and the rural birth rate was 23.8. The superior position of the towns is doubtless due to a more favorable environment, but it may also be in part due to the fact that the parents of these scientific men were the abler clergymen and others of their generation who were drawn to the cities." ['06, pp. 738-739]

TABLE 37.

Distribution in Different Places.

	According to Birthplace			Per Million
	I.-V.	VI.-X.	Total	1860.
New York, N. Y.	33	25	58	71.2
Boston, Mass.	24	19	43	241.8
Philadelphia, Pa.	12	16	28	49.5
Baltimore, Md.	9	11	20	94.1
Cincinnati, O.	6	6	12	74.5
Brooklyn, N. Y.	3	8	11	39.4
Chicago, Ill.	5	3	8	73.2
Buffalo, N. Y.	3	4	7	86.2
St. Louis, Mo.	2	5	7	43.5
Cambridge, Mass.	4	2	6	230.2
Cleveland, O.	4	2	6	140.5
Salem, Mass.	1	5	6	269.6
Milwaukee, Wis.	1	4	5	110.5
Newark, N. J.	3	2	5	69.5
San Francisco, Cal.	2	3	5	88.0
Total	112	115	227	

With respect to the circumstances of education "it appears that those who attend the larger universities are not of higher average performance than others. . . . . .

"There is no significant difference in rank between the 515 men who attended the larger institutions and those who attended smaller colleges or none. It might be supposed that abler students would be attracted to a university such as Harvard, and that they would have greater opportunities there, but this appears not to be the case. So far as it goes, this favors the theory that men of science are born such and are not dependent on the environment for the quality of their performance. . . . .

"The conditions are similar in the case of the doctor's degree." ['06, pp. 740-741]

Dr. Rice's study is quoted at some length because it was the first of a series of studies of the actual results of school work, still few in number, but destined to increase rapidly with increasing scientific interest in school administration.

Dr. Rice ['97] tested the spelling ability of some 33,000 children in twenty-one schools representing a great variety in spirit, methods, time given to spelling and in other respects. He then compared the conditions in schools where the pupils did well in spelling with those in schools where they did badly. He notes first of all the slight differences between schools, only 6 out of the 21 schools being outside the limits 73.3 and 77.9, and the decrease in variation amongst schools as we pass from lower to higher grades (see Table 38), facts which show that the differences in spirit or method that characterize schools can not make much difference in achievement. Of school systems where mechanical methods are in use as compared with more progressive systems he says:

"Indeed, in both the mechanical and the progressive schools the results were variable; so that while, in some instances, the higher figures were secured by the former, in others they were obtained by the latter; and the same is true of the lower figures. For example, School B, No. 11, in which the best average (79.4) was obtained, belongs to a very progressive system; while School A, No. 12, which made only 73.9, belongs to one of our most mechanical systems. And it is a peculiar incident that, in both these cities, the results in the only other school examined are exactly reversed, although the environment is about the same."

He eliminates the possibility that home reading or cultured parents or English rather than foreign parentage is the cause of the differences amongst schools by making the comparisons of Table 39.

Dr. Rice further tabulated the results in accordance with the methods of instruction used in the different schools, interviewing some two hundred teachers for that purpose. He does not give the detailed results, but assures us that there is no reason to believe that there is any clear choice between oral and written spelling, writing isolated words and writing sentences, the sight or flash method and its absence. Phonic reading does not make bad spellers, nor do written language work and wide general reading make good spellers. "In brief," says he, "there is no direct relation between method and results. . . . The results varied as much under the same as they did under different methods of instruction."

TABLE 38.

AVERAGES FOR INDIVIDUAL SCHOOLS IN SPELLING.

City	School	4th Year Sentence-Test	4th Year Composition-Test	4th Year Minutes Daily	5th Year Sentence-Test	5th Year Composition-Test	5th Year Minutes Daily	6th Year Sentence-Test	6th Year Composition-Test	6th Year Minutes Daily	7th Year Sentence-Test	7th Year Composition-Test	7th Year Minutes Daily	8th Year Sentence-Test	8th Year Composition-Test	8th Year Minutes Daily	School Av., Sentence-Test	City	School
1	A	67.6	—	—	79.6	—	45	77.2	—	20	76.7	—	25	—	—	—	77.1	1	A
1	B	—	—	—	—	—	—	—	—	—	—	—	—	82	—	—		1	B
1	C	61.8	96.8	50	72.6	97.4	50	77.7	98.8	50	—	—	40	86.1	99.6	20	75	1	C
		71	96.7	45	81.8	98.4	35	71.5	98.2	40	78	98.7	30	85.8	99	30	77.6		
7	A	B 77	95.9	30	73.6	97.8	20	64.2	97.8	20	73.1	—	—	84.9	—	30		7	A
		A 68.6	97.5	40	79.4	98.3	15	73.4	98.6	40	—	99.4	40	—	99.4	35	75.2		
7	B	B 64.5	97	—	75	—	30	71.1	—	45	77.2	—	30	87.2	99.4	—		7	B
		A 75.8	97.9	20	81.1	—	40	80	—	50	81.1	—	15	86.7	99.3	30	77.9		
7	C	B —	97.5	—	—	97.6	—	—	97.8	35	—	98.7	—	—	99.4	15		7	C
		A —	97.4	35	—	98.6	15	—	99.1	45	—	99.2	30	—	—	—	—		
9	A	B 70.8	—	—	71.3	—	—	66.5	—	—	85.6	—	—	—	—	—		9	A
		A 66.6	—	40	78	—	—	72.8	—	35	—	—	50	—	—	—	76.6		
9	B	B 66.4	96.1	—	74.8	97.6	—	73.2	98.7	35	—	99.4	—	—	—	—		9	B
		A 68.4	96.8	—	76.8	97.9	45	75.1	98.4	45	86.5	99.2	60	—	—	—	77.7		
9	C	B 66	—	—	70	—	—	76.1	—	40	—	—	—	—	—	—		9	C
		A 66.4	—	40	83.2	—	—	79.4	—	35	84.7	—	20	—	—	—	77.9		
10	A	B —	—	—	—	—	—	—	—	—	80	99.2	—	83.5	99.2	—		10	A
		A —	—	—	—	—	—	—	—	—	79.5	98.6	20	82.5	99.1	20	76.2		
10	B	B 61.4	96.6	—	76.6	98.6	—	72.7	99.2	—	—	—	—	—	—	—		10	B
		A 75.9	98.3	35	77.2	98.9	20	—	—	20	—	—	—	—	—	—	—		
11	A	B 61.2	—	—	70.4	—	—	63.4	—	—	—	—	—	—	—	—		11	A
		A 63.6	—	25	70.6	—	—	65.8	—	25	76	—	30	86.4	—	40	72		

Small B indicates first half, and small A second half of school year.

TABLE 38. (Continued).

AVERAGES FOR INDIVIDUAL SCHOOLS IN SPELLING.

City	School		4th Year Sentence-Test	4th Year Composition-Test	4th Year Minutes Daily	5th Year Sentence-Test	5th Year Composition-Test	5th Year Minutes Daily	6th Year Sentence-Test	6th Year Composition-Test	6th Year Minutes Daily	7th Year Sentence-Test	7th Year Composition-Test	7th Year Minutes Daily	8th Year Sentence-Test	8th Year Composition-Test	8th Year Minutes Daily	School Av., Sentence-Test	City	School
11	B	B	76.4	96.3	—	74.6	97.9	—	74	98.5	—	—	—	—	—	—	—		11	B
		A	79.8	97.9	35	81.8	98.5	25	—	—	30	80	98.8	40	86.7	99.1	40	79.4		
12	A	B	53.2	—	—	66.4	—	—	75	—	—	—	—	—	86.3	—	—		12	A
		A	63.1	—	—	73	—	—	75.7	—	—	78.3	—	—	90.2	—	—	73.9		
12	B	B	53.3	—	—	74.4	—	18	73.1	—	—	—	—	—	90.3	—	—		12	B
		A	67.2	—	18	79	—	10	78.4	—	9	90	—	12	94.6	—	12	79.0		
15	B	B	—	—	—	73	96.8	—	57	97.7	—	76.7	98.3	—	83.9	99	—		15	B
		A	—	—	—	73.4	97.4	—	71.5	98.1	—	83.3	93.7	—	87.2	99.4	—	73.4		
15	D	B	—	—	—	81.6	—	—	68.3	—	—	74.7	—	—	86.2	—	—		15	D
		A	74.3	—	20	83.6	—	20	72.7	—	20	84	—	20	88.6	—	20	78.8		
15	B	B	66.2	97	—	74.4	97.8	—	72.8	98.5	—	—	—	—	—	—	—		15	E
		A	70.8	96.6	15	79.4	98.3	20	69.6	98.1	20	80.6	98.7	20	89.4	99.2	20	77.3		
15	H	B	67.2	—	—	76.2	—	—	69.8	—	—	81.4	—	—	89.9	—	—		15	H
		A	69.1	—	20	79.1	—	25	76.2	—	25	—	—	25	89.4	—	25	77.9		
16	A	B	62.6	97.4	—	73.6	98.5	—	61.9	97.6	—	71.1	93.2	—	84	99.1	—		16	A
		A	59.2	97.9	20	76.4	98.6	15	68.7	98.7	20	72.8	98.8	20	89.9	99.4	6	72.0		
16	B	B	68.4	—	—	70.6	98.5	—	—	98.1	—	71.3	98.5	—	81.2	98.9	—		16	B
		A	65	97.9	20	74.4	98.2	20	—	97.7	15	75.3	99.3	10	80.2	99.2	5	72.7		
19	A	B	61.8	—	—	66.8	—	—	68.8	—	—	78	—	—	84.3	—	—		19	A
		A	—	—	20	76.6	—	40	72	—	35	84	—	—	—	—	—	73.8		
19	E	B	57.4	—	—	73	—	—	61.3	—	—	69.5	—	—	85	—	—		19	B
		A	68.4	—	40	78.2	—	—	69.5	—	30	84	—	30	86.8	—	—	73.3		

Small B indicates first half, and small A second half of school year.

TABLE 39. (Table 3 of the original account)

	Grade	No. of Cities	No. of Classes	No. of Pupils	General Average	Children of Foreign Parentage	Average	No. of Children Hearing Foreign Language at Home	Average	Children of Unskilled Laborers	Average
Sentence Test	Fourth ..	4	27	821	64.7	155	65.2	159	64.9	129	62.5
	Fifth....	4	29	829	76.	153	77.4	157	76.7	129	74.5
	Sixth....	4	22	778	69.7	185	69.6	165	70.3	119	70.4
	Seventh .	4	18	566	78.8	81	82.5	52	81.5	55	76.8
	Eighth...	4	19	528	83.1	72	83.2	64	83.2	76	85.

That the amount of time given is not the cause of success in teaching spelling is shown by the facts of Table 38. Schools giving 15 or 20 minutes daily to spelling do as well as those giving 40 or 50.

After this admirable array of facts Dr. Rice jumps rather hastily to this speculative conclusion: "The facts here presented, in my opinion, will admit of only one conclusion, viz., that the results are not determined by the methods employed, but by the ability of those who use them. In other words, the first place must be given to the personal equation of the teacher, while methods and devices play a subordinate part."

This statement should have been based upon a demonstration of a high coefficient of correlation between the measure of a class in spelling and the measure of its teacher in ability, or of a great increase in variability in spelling ability as we pass from the children taught by one teacher to the children taught by 10 or 20 different teachers. I calculate that if the reliabilities of Dr. Rice's eighth grade averages are what they would seem to be from tests made in eighth grades by myself and my students,* the differences amongst them are not much greater than we would expect by the law of chance if the teaching were in all cases equally efficient. The average deviation from their mean of the 12 eighth grade classes which were tested in the first half of the year is 1.9; that of the 13 tested in the last

*These give a variability of 12.2 amongst the individuals of the grade.

half of the year is 2.6; the average deviation by pure chance of 12 eighth grade classes of 40 students each would be 1.9, the variability of individuals being 12.2. So, in the case of the eighth grades, we may need no cause at all for the differences amongst schools save the inaccuracy of the averages due to the small number of cases.

Dr. Rice measured the arithmetical ability of some 6,000 children in 18 different schools in 7 different cities ['02]. The results of these measurements are summarized in Table 40. This table "gives two averages for each grade as well as for each school as a whole. Thus, the school at the top shows averages of 80.0 and 83.1, and the one at the bottom, 25.3 and 31.5. The first represents the percentage of answers which were absolutely correct; the second shows what per cent. of the problems were correct in principle, *i. e.*, the average that would have been received if no mechanical errors had been made. The difference represents the percentage of mechanical errors, which, I believe, in most instances, makes a surprisingly small appearance."

From these results Dr. Rice seeks the causes of excellence in arithmetical work, as in the case of spelling, by comparing the conditions in the successful schools with those in the unsuccessful. He deals *seriatim* with (1) the home environment of the pupils; (2) the size of the classes; (3) the age of the children; (4) the time of day of the test; (5) the time devoted to arithmetic in the school; (6) the amount of home work required; (7) the methods of teaching; (8) teaching ability as represented by a combination of education, training and the personality of the teacher; (9) the course of study; (10) the superintendent's training of teachers; (11) the establishment of demands in regard to results; (12) the testing for results (*a*) by teachers alone, (*b*) by teachers and superintendents, (*c*) by principals, (*d*) by principals and superintendents.

He finds that the work depends upon the method of testing for results, that teachers and pupils do about what is demanded of them, and that the best work appears when the superin-

TABLE 40.

AVERAGES FOR INDIVIDUAL SCHOOLS IN ARITHMETIC.

	School	4th Year		5th Year		6th Year		7th Year		8th Year		School Average			Minutes Daily	When Taken
		Result	Principle	Result	Principle	Result	Principle	Result	Principle	Result	Principle	Result	Principle	Per Cent of Mechanical Errors		
City III	1	68.4	76.7	79.5	82.5	79.3	80.3	81.1	82.3	91.7	93.9	80.0	83.1	3.7	53	A.M.
City I	1	72.7	81.4	84.7	88.8	80.4	81.5	64.2	67.2	80.9	82.8	76.6	80.3	4.6	60	A.M.
City I	2	......	......	80.3	87.1	80.9	83.4	43.5	50.9	72.7	79.1	69.3	75.1	7.7	25	A.M.
City I	3	54.5	66.0	74.7	78.3	72.2	74.0	63.5	66.2	74.5	76.6	67.8	72.2	6.1	45	A.M.
City I	4	60.0	73.3	70.8	79.3	69.6	72.2	54.6	57.8	66.5	69.1	64.3	70.3	8.5	45	P.M.
City II	1	81.3	87.7	78.2	85.6	71.2	75.3	33.6	35.7	36.8	40.0	60.2	64.8	7.1	60	P.M.
City III	2	70.1	79.9	53.6	60.0	43.7	45.0	53.9	56.7	51.1	53.1	54.5	58.9	7.4	60	P.M.
City IV	1	70.5	76.4	73.2	77.7	58.9	60.4	31.2	34.1	41.6	43.5	55.1	58.4	5.6	60	A.M.
City IV	2	62.9	72.8	70.5	76.8	59.8	63.1	......	......	22.5	22.5	53.9	58.8	8.3	..	P.M.
City IV	3	59.8	72.7	65.3	73.5	54.9	58.1	35.2	38.6	43.5	45.0	51.5	57.6	10.5	60	A.M.
City IV	4	53.5	62.2	53.5	65.7	42.3	45.1	16.1	19.2	48.7	48.7	42.8	48.2	11.2	..	P.M.
City V	1	38.5	46.3	67.0	71.0	44.1	48.7	29.2	32.5	51.1	58.3	45.9	51.3	10.5	40	P.M.
City VI	1	28.1	31.7	38.1	44.2	68.3	71.3	33.5	36.6	26.9	30.7	39.0	42.9	9.0	33	A.M.
City VI	2	41.6	51.9	45.3	52.1	46.1	49.5	19.5	24.2	30.2	40.6	36.5	43.6	16.2	30	P.M.
City VI	3	36.8	54.2	55.0	62.8	34.5	36.4	30.5	35.1	23.3	24.1	36.0	42.5	15.2	48	P.M.
City VII	1	59.3	69.3	53.7	63.1	35.2	37.7	29.1	32.5	25.1	27.2	40.5	45.9	11.7	42	A.M.
City VII	2	47.4	55.2	65.4	71.4	35.2	38.7	15.0	16.4	19.6	21.2	36.5	40.6	10.1	75	P.M.
City VII	3	41.1	58.0	37.5	44.5	27.6	33.7	8.9	10.1	11.3	11.3	25.3	31.5	19.6	45+	P.M.
General average	...	59.5	69.9	69.4	75.5	60.7	63.2	39.4	42.5	49.4	51.9	55.7	60.6	8.1		
Per cent of mechanical errors	...	......	14.8	......	8.1	......	3.9	......	7.3	......	4.8					
Number of pupils examined	...	......	1,422	......	1,593	......	1,285	......	974	......	689	Total	5,963			

tendent, in connection with principals of schools, tests and rates the work of the classes.

The following are samples of the reasoning by which he eliminates one after another of the possible causes:

Home Environment

"If the part that is played by the home environment should be as important as it is generally supposed to be, we should, of course, expect to find that the schools represented in the upper part of the table had been attended by children from cultured homes, while those in the lower part had been attended by those whose home environment was very poor. However, if a line should be drawn across the middle of the table, and the schools above it compared with those below, such a condition would not be found. Indeed, careful inspection would show that the odds were certainly not in favor of the 'aristocratic' districts. Of the eighteen schools, three in particular are representative of the latter, and the best of these secured the tenth place, while the others ranked eleventh and sixteenth, respectively. The school that ranked seventh was distinctively a school of the slums. That is to say, the school laboring under the poorest conditions in respect to home environment obtained a better standing than any one of the so-called aristocratic schools. The building which stands fifth is representative of conditions just a shade better than those of the slums. And when I add that, from the standpoint of environment, the schools of City I. did not average a single degree better than those of Cities VI. and VII., I have said enough to show that the poor results secured in the latter cities can not be condoned on the ground of unfavorable environment. Thus, as in spelling, so in arithmetic, this mountain, upon close inspection, dwindles down to the size of a molehill."

Size of Classes

"Equally surprising, if indeed not more incredible, may appear the statement that no allowance whatever is to be made for the size of the class in judging the results of my test. I shall not enter into the details in regard to this point, but will dismiss it with the remark that the number of pupils per class was larger in the highest six schools than it was in the schools of City VI., and that the classes were exceptionally small in the school that stands at the lower end."

Age of Pupils

His argument is here too lengthy to quote and is rather awkward, but sufficiently proves that the differences between schools could have been due only in a very slight degree, if at all, to differences in the ages of the pupils. The obvious way to eliminate age is to compare a group from City VI. or VII. with a group identical in age and grade from City I. or III.

Time of Day

This can not be the cause of much of the difference found, for within any one city the time of day of the test makes little difference.

The Time Devoted to Arithmetic in the School

"A glance at the figures will tell us at once that there is no direct relation between time and result; that special pressure does not necessarily lead to success, and, conversely, that lack of pressure does not necessarily mean failure.

"In the first place, it is interesting to note that the amount of time devoted to arithmetic in the school that obtained the lowest average—25 per cent.—was practically the same as it was in the one where the highest average—80 per cent.—was obtained. In the former the regular time for arithmetic in all the grades was forty-five minutes a day, but some additional time was given. In the latter the time varied in the different classes, but averaged fifty-three minutes daily. This shows an extreme variation in results under the same appropriation of time.

"Looking again toward the bottom of the list, we find three schools with an average of 36 per cent. In one of these, insufficient pressure might be suggested as a reason for the unsatisfactory results, only thirty minutes daily having been devoted to arithmetic. The second school, however, gave forty-eight, while the third gave seventy-five. This certainly seems to indicate that a radical defect in the quality of instruction can not be offset by an increase in quantity.

"If we now turn our attention from the three schools just mentioned and direct it to three near the top—Schools 2, 3 and 4, City I.—we find the conditions reversed; for while the two schools that gave forty-five minutes made averages of 64 per cent. and 67 per cent., respectively, the school that gave only twenty-five minutes succeeded in obtaining an average of 69 per cent. This would appear to indicate that while, on the one hand, nothing is gained by an increase of time where the instruction in arithmetic is faulty, on the other hand, nothing is lost by a decrease of time, to a certain point, where the schools are on the right path in teaching the subject. Perhaps the most interesting feature of the table is the fact that the school giving twenty-five minutes a day came out within two of the top, while the school giving seventy-five minutes daily came out practically within one of the bottom."

The Amount of Home Work Required

The greatest amount of home work was required in the lowest ranking city while it had been practically abandoned in the first five schools of the table.

In the other cases the facts are given more vaguely, and in his presentation of positive evidence that differences in supervision by tests are the leading causes of the differences in achievement of the different schools, Dr. Rice seems to reach his conclusion simply from observing (1) that all conditions in City VI. were favorable save that examinations were given only by the teacher, (2) that in City VII. the examinations were given by the teacher and perfunctorily by the superintendent, while (3) in City I. the superintendent, with the principals, took pains in setting the tests. It seems probable that the cause he alleges is a real one, though even his own facts show the co-operation of other causes. This I take it he does not mean to deny.

The fourth sample of studies of the influence of the environment is not of major importance, but is distinctive in that its facts cannot be accounted for by any force other than the environment. The facts are the changes in the careers of scholarly college graduates from the class of 1840 to that of 1895, comprising 5283 members of the honorary society, Φ B K, admission to which was substantially a recognition of superior scholarship in college.

The four professions, law, medicine, teaching and the ministry, have, together, attracted almost exactly the same proportion of scholarly men in each decade. The per cent of Φ Β Κ graduates entering some one of these four professions was 65 in 1840-59, 65½ in 1860-69, 65 in 1870-79, and 64 in 1880-1894.

Among the professions, however, there have been marked changes, as shown in Table 41. In twenty years the law doubled its attractiveness to scholarly men and then, in half that time, lost two-thirds of its gain. Medicine was, in the last decade of the period, becoming more attractive. The table shows a very rapid rise in the popularity of teaching from 1840 to 1860 and again from 1870 to 1895. The years from '60 to '65 show an opposite tendency. The law was then gaining rapidly and the ministry was holding its own. The most striking change was the decrease in the proportion of scholarly men making the ministry their life work. The decrease would be even more marked if those who entered the ministry but gave up its regular work for that of teaching were included. The incomplete records available in the Φ Β Κ Catalogue of 1900 give only 5½ per cent of clergymen amongst those graduating from '95-'99; and even with later additions the per cent for 1900 is probably under 10.

Roughly, it may be said that three-fourths of the scholarly young men who entered the ministry in 1850 would have gone into teaching or the law if they had happened to be born a half century later. The same original natures choose differently because the social and intellectual environment has changed.

The near future will doubtless see a rapid increase in the number and improvement in the quality of studies of the environmental causes of individual differences in mental traits. Rice's investigation of the differences due to different features of administration and teaching has been followed by similar studies by Cornman ['02], Stone ['08], Courtis ['09 and later] and Thorndike ['10]. Experts in education are becoming experimentalists and quantitative thinkers and are seeking to verify or refute the established beliefs concerning the effects of

TABLE 41.

PERCENTAGES OF SCHOLARLY YOUTHS MAKING THEIR LIFE WORK THAT OF:—

	Law	Medicine	Teaching	Ministry
1840—1844	14	6 (1840—1859)	9.4	37.5
1845—1846	14		11.6	40
1850—1854	9.3		13.7	36.5
1855—1859	10.5		16.4	34.5
1860—1864	15.2	5.5	17.2	27.5
1865—1869	19.7	4	13.9	28.5
1870—1874	19.8	5.5	16.4	22.5
1875—1879	22.5	4	17.6	22
1880—1884	16.4	4.5	21.4	19.5
1885—1889	14.4	7.5	25.5	16
1890—1894	19	7	25.4	14

educational forces upon human nature. Students of history, government, sociology, economics, ethics and religion are becoming, or will soon become, quantitative thinkers concerning the shares of the various physical and social forces in making individual men differ in politics, crime, wealth, service, idealism, or whatever trait concerns man's welfare.

To the facts presented in these four sample studies and in previous chapters, we may add certain very significant measures of the effect of equal amounts of exercise of a function upon individual differences in respect to efficiency in it. The argument is as follows: In so far as the differences in achievement found amongst a group of men are due to differences in the quantity and quality of training which they have had in the function in question, the provision of equal amounts of the same sort of training for all individuals in the group should act to *reduce* the differences. Suppose, for example, that eleven individuals showed efficiencies of 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 and 20 respectively in the number of words that they could typewrite per minute. Suppose that this variation had been entirely caused by a corresponding variation in the amount of time they had spent in practicing typewriting, say 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15 hours. Then giving each individual 10 hours more of practice, so that the range in respect to amount of practice would be from 15 to 25 hours, should result in reducing the relative differences. The person who now had had 15 hours of practice should by the hypothesis show an efficiency of 20, while the person with 25 hours of practice would not be expected to be beyond 30. Whereas the limiting scores were as 2 to 1 (20 to 10), they should not now differ more than as 3 to 2 (30 to 20).*

If the addition of equal amounts of practice does *not* reduce the differences found amongst men, those differences can not well be explained to any large extent by supposing them to have been due to corresponding differences in amount of previous practice. If, that is, inequalities in achievement are not reduced by equalizing practice, they cannot well have been caused by inequalities in previous practice. If differences in opportunity cause the differences men display, making opportunity more nearly equal for all by adding equal amounts to it in each case should make the differences less.

The facts found are rather startling. Equalizing practice *seems to increase differences.* The superior man seems to have got his present superiority by his own nature rather than by superior advantages of the past, since, during a period of equal advantages for all, he increases his lead.

The available measurements of the effect of adding equal amounts of practice in the case of individuals of different ability are those of Thorndike ['08a and '10a], Whitley ['11], Starch ['11], Wells ['12], Kirby ['13], Donovan and Thorndike ['13], and Hahn and Thorndike ['14]. I quote samples of the results obtained.

The following table (Table 42) giving the initial and final scores in practice at the mental multiplication of a three-place by a three-place number, speaks for itself. The same effect appears, though less emphatically, in the case of Whitley's nine individuals in a similar experiment ['11]. The four who were

*The exact expectation would, of course, depend upon the form of the practice-curve in question for the function in question and the cooperating factors in the learners; the illustration is made arbitrarily simple.

TABLE 42.

THE EFFECT OF EQUAL AMOUNTS OF PRACTICE UPON INDIVIDUAL DIFFERENCES IN THE MENTAL MULTIPLICATION OF A THREE-PLACE BY A THREE-PLACE NUMBER.

		Amount done per unit of time.			Percentage of correct figures in answers.		
	Hours of Practice	First 5 Examples	Last 5 or 10 Examples	Gain	First 5 Examples	Last 5 or 10 Examples	Gain
Initial highest five individuals	5.1	85	147	61	70	78	18
" next five "	5.1	56	107	51	68	78	10
" " six "	5.3	46	68	22	74	82	8
" " six "	5.4	38	46	8	58	70	12
" " five "	5.2	31	57	26	47	67	20
" " one individual	5.2	19	32	13	100	82	*-18*

most efficient at the start made a greater average gain from equal practice than the four who were least efficient.

Using the multiplication of a three-place by a one-place number Starch ['11] got results showing the same effect. Of his eight subjects, the three best averaged 39 examples per 10 minutes in the initial test and gained on the average 45 in the course of doing 700 examples. The three lowest, who averaged 25 examples per 10 minutes in the initial test, gained only 26 in the course of doing 700 examples, in spite of the fact that 700 examples represented for them a much greater amount of practice, measured in time spent.

In the case of addition [Thorndike, '10], the initially highest individuals of nineteen adults showed this same greater gain in amount done at equal accuracy per unit of time when all nineteen were given approximately equal practice. The facts are given in Table 43. Similar results have been obtained by Wells ['12], Kirby ['13], Hahn and Thorndike ['14] and others.

These experiments concerning the effect of practice upon individual differences in mental multiplication, addition, mark-

TABLE 43.

THE EFFECT OF EQUAL AMOUNTS OF PRACTICE UPON INDIVIDUAL DIFFERENCES IN COLUMN-ADDITION OF ONE-PLACE NUMBERS.

	Average number of additons per 5 minutes corrected for errors			Average time spent in practice from mid-point of first test to mid-point of last test (in minutes)
	First test	Last test	Gain	
Initially highest 6 individuals	297	437	140	40
Initially next highest 6 individuals	234	345	111	49
Initially lowest 7 individuals	167—	220+	54	46

ing A's on printed sheets of capitals and the like are too restricted in scope and in the amount of practice to justify any general application of their results. In other mental functions the achievements of a man in comparison to his fellows may be more a consequence of his advantages and less a consequence of his own nature. So far as they go, however, experiments in practice have given no support to the common assumption that differences in external conditions are responsible for the bulk of the variation found among men of the same race and general social status.

THE METHOD OF ACTION OF DIFFERENCES IN ENVIRONMENT

We may summarize the methods whereby different environments act upon intellect and morals as:—

1. Furnishing or withholding the physiological conditions for the brain's growth and health.

2. Furnishing or withholding adequate stimuli to arouse the action of which the brain is by original nature or previous action capable.

3. Reinforcing some and eliminating others of these activities in consequence of the general law of effect.*

*In all animals capable of profiting by training any act which in a given situation brings satisfaction becomes thereby more closely associated with

According to this description we should look upon the mental life of an individual as developing in the same way that the animal or plant kingdom has developed. As conditions of heat and food-supply have everywhere been the first requisite to and influence on animal life, so the physiological conditions of the brain's activities are the first modifiers of feeling and action. As the stimuli of climate, food, unknown chemical and electrical forces and the rest have been the means of creating variations in the germs or of stimulating to action the inner tendency of the germs to vary, and so have rendered possible the production of millions of different animal types, so the sights and sounds and smells of things, the words and looks and acts of men, the utensils and machinery and buildings of civilization, its pictures and music and books, awaken in the mind new mental varieties, new species of thoughts and acts. In a score of years from birth the human mind, like the animal world, originates its universe of mental forms. And as, in the animal kingdom, many of these variations fail to fit the conditions of physical nature and die after a generation or two, so in any one of us many of the mental forms produced are doomed to a speedy disappearance in consequence of their failure to fit outside events. The elimination of one species by others in the animal world is again paralleled by the death of those thoughts or acts which are out of harmony with others. Species of thoughts, like species of animals, prey upon one another in a struggle in which survival is the victor's reward. Further, just as species of animals fitted to one environment perish or become transformed when that environment changes, so mental forms fitted to infancy perish or are transformed in school life; mental forms fitted to school life perish in the environment of the workaday world; and so throughout the incessant changes of a mind's surroundings. In mental life resulting pain or discomfort is the cause of the extinction of a species. The condi-

that situation, so that when that situation recurs the act will recur also. An act that brings discomfort becomes dissociated from the situation and less likely to recur.

tion of a man's mind at any stage in its history is then, like the condition of the animal kingdom at any stage in the history of the world, the result not only of the new varieties that have appeared, but also of a natural selection working upon them. The tale of a human mind's progress is the tale of the extinction of its failures. Possibility of existence, stimuli to variations, selection by elimination: these words that describe the action of the environment on animal life are equally competent to tell the record of a human life.

The influence of any environmental agency, physical or social, varies with its avoidability. Oligarchies lose in influence if there is a democracy to which men may emigrate. Customs do not make men so infallibly if there is a radical party, however small, which offers an alternative mode of life. Music's charms to soothe obviously are not so universal if men can close their ears. A creed loses authority as soon as one disbeliever seeks converts. Social environments, institutions, beliefs and modes of behavior are nearly omnipotent when undisputed; for to be the first man to revolt means either that one is a mere eccentric and so sure to be a failure, or that one is a genius and so very rare. But once a revolt is started and advertised, it may much more easily attract those whose original natures it fits. And they may be the more attracted by it for having been exposed to the opposite force. So a given environmental force may even act as a stimulus toward just the opinions, interests or acts that it is designed to thwart.

There are many differences in thought and conduct which are nearly equally tolerated by all original natures. To wear a hat or not to wear a hat, to express requests and opinions in English or to express them in German, to learn astrology or to learn the Ptolemaic astronomy or to learn the Copernican astronomy—to all original natures these are nearly indifferent issues. Which is done depends almost exclusively on environment. In general this is true of all the 'whats' of knowledge and technique. *How many* and *how hard* things a man can learn or do are largely decided by original nature, but, within these limits, *what* he learns or does is largely a matter of what he is stimulated to do and rewarded for doing. On the other hand, there are many features of original nature each of which acts to produce nearly the same effect in spite of such differences in outside forces as different men can meet in modern civilized countries. In such countries it seems possible for any one to be a poet, or to be a political leader, or to be a money-maker, if his nature so orders. Original nature in general is not irrepressible, and no form of it is absolutely irrepressible; but some forms of original nature seem to be nearly irrepressible by any of the environments a man in this country today is likely to have.

THE RELATIVE IMPORTANCE OF ORIGINAL NATURE AND ENVIRONMENT

It is impossible at present to estimate with security the relative shares of original nature, due to sex, race, ancestry and accidental variation, and of the environment, physical and social, in causing the differences found in men. One can only learn the facts, interpret them with as little bias as possible, and try to secure more facts. This interpretation is left to the student, but with certain cautions in addition to or in amplification of those already explained.

Many of the false inferences about nature *versus* nurture are due to neglect of the obvious facts:—that if the environments are alike with respect to a trait, the differences in respect to it are due entirely to original nature; that if the original natures are alike with respect to a trait, the differences in respect to it are due entirely to differences in training; and that the problem of relative shares, where both are effective, includes all the separate problems of each kind of environment acting with each kind of nature. Any one estimate for all cases would be absurd.

Many disagreements spring from a confusion of what may be called absolute achievement with what may be called relative achievement. A man may move up a long distance from zero and nevertheless be lower down than before in comparison with other men: absolute gain may be relative loss. One thinker may attribute differences in achievement almost wholly to nurture while another holds nature to be nearly supreme, though both thinkers possess just the same data, if the former is thinking of absolute and the latter of relative achievement. The commonest error resulting is that of concluding from the importance of sex and ancestral heredity that education and social control in general are futile. On the contrary, as I have elsewhere said, such studies as those of Chapters IX, X, and XI merely prove the existence of, and measure certain determinants of, human intellect and character and demonstrate that the influences of the environment are differential, the product varying not only in accord with the environmental force itself but also in accord with the original nature upon which it operates. We may even expect that education will be doubly effective, once society recognizes the advantages given to some and denied to others by heredity. That men have different amounts of capacity does not imply any the less advantage from or need of wise investment. If it be true, for example, that the negro is by nature unintellectual and joyous, this does not imply that he may not be made more intelligent by wiser training or misanthropic and ugly-tempered by the treatment he now receives. It does mean that we should be stupid to expect the same results from him that we should from an especially intellectual race like the Jews, and that he will stand with equanimity a degree of disdain which a Celt would requite with dynamite and arson.

To the real work of man for man,—the increase of achievement through the improvement of the environment,—the influence of heredity offers no barrier. But to the popular demands from education and social reforms it does. For the common man does not much appreciate absolute happiness or absolute betterment. He does not rejoice that he and his children are healthier, happier and more supplied with noble pleasures than were his ancestors of a thousand years ago. His complaint is that he is not so well off as some of those about him; his pride is that he is above the common herd. The common man demands *relative* superiority,—to be above those of his own time and locality. If his son leads the community, he does not mind his real stupidity; to be the handsomest girl in the county is beauty enough. Social discontent comes from the knowledge or fancy that one is below others in welfare. The effort of children in school, of men in labor and of women in the home is, except as guided by the wise instincts of nature or more rarely by the wisdom of abstract thought, to rise above some one who seems higher. Thus the prizes which most men really seek are after all in large measure given or withheld by original nature. In the actual race of life, which is not to get ahead, but to get ahead of somebody, the chief determining factor is heredity.

But the prizes which education *ought* to seek are all within its power. The results for which a rational mankind would strive are determined largely by mankind itself. For the common good it is indifferent *who* is at the top,—*which* men are achieving most. The important thing for the common good, for all men, is that the top should be high—that much should be achieved. To the absolute welfare of all men together education is the great contributor.

Another caution is not to make false inferences about moral responsibility from the fact that individual differences are in large measure due to nature; nor to use such false inferences to discourage acceptance of evidence in support of this fact.

It is from time to time complained that a doctrine which refers mental traits largely to original make-up, and consequently to ancestry, discourages the ambitions of the well-intentioned and relieves the world's failures from merited contempt. But every one is agreed that a man's free will works only within limits, and it will not much matter for our practical attitude whether those limits are somewhat contracted. If the question is between original nature and the circumstances of nurture it is rather more encouraging to believe that success will depend on inherent qualities than to refer it entirely to advantages possessed during life, and contempt is merited more by him who has failed through being the inferior person than by the one who has failed simply from bad luck. Whether or not it is merited in either of the two cases we shall decide in view of our general notions about merit and blame, not of our psychological theories of the causes of conduct.

On the whole it seems certain that prevalent opinions much exaggerate the influence of differences in circumstances and training in producing the intellectual and moral differences found in men of the same nation and epoch. Certain natures seem to have been made by certain environments when really the nature already made selected that environment. Certain environments seem to eliminate certain traits from an individual when really they merely expel the individual *in toto*.

Thinkers about the organized educational work of church, library and school need especially to remember three facts.

First.—For the more primitive and fundamental traits in human nature such as energy, capability, persistence, leadership, sympathy and nobility the whole world affords the stimulus, a stimulus that is present well-nigh everywhere. If a man's original nature will not respond to the need of these qualities and the rewards always ready for them, it is vain to expect much from the paltry exercises of the schoolroom.

Second.—The channels in which human energy shall proceed, the specific intellectual and moral activities that shall profit by human capacities, are less determined by inborn traits. The schools should invest in profitable enterprises the capital nature provides. We can not create intellect, but we can prevent such a lamentable waste of it as was caused by scholasticism. We can not double the fund of human sympathy, but we can keep it clear of sentimental charity.

Third.—Morality is more susceptible than intellect to environmental influence. Moral traits are more often matters of the direction of capacities and the creation of desires and aversions. Over them then education has greater sway, though school education, because of the peculiar narrowness of the life of the schoolroom, has so far done little for any save the semi-intellectual virtues.

The one thing that educational theorists of today seem to place as the foremost duty of the schools—the development of powers and capacities—is the one thing that the schools or any other educational forces can do least. The one thing that they can do best is to establish those particular connections with ideas which we call knowledge and those particular connections with acts which we call habits.

CHAPTER XIV

THE NATURE AND AMOUNT OF INDIVIDUAL DIFFERENCES IN SINGLE TRAITS

For the purpose of the following discussion, let a 'single trait' be defined as one whose varying conditions in men can be measured on one scale. A combination of traits requires two or more scales. For example, in so far as the difference between John and James in reaction time to sound can be measured as so many thousandths of a second on one scale, reaction time to sound is a single trait. The difference between John and James in temperament, on the contrary, can be stated only in terms of several scales, such as quick . . . . . slow, intense . . . . . superficial, broad . . . . . narrow, and the like. So temperament is to be regarded as a combination of traits.

Individuals may be compared with respect to one trait at a time, or with respect to certain combinations of traits. We naturally take up first the simpler case.

The most desirable description of the differences between individuals in a mental trait would be to give the facts for all human beings,* then for all of each sex;* then for all of each race and of one sex,* and so on for each stock, degree of maturity, and kind of training. That is, the condition of the individuals in each of a great many groups, each defined by sex, ancestry, age and training, would be described. By combining these one could describe the condition of the individuals in groups defined by sex and age alone, or race and age alone, or race, age and training alone, and the like.

For practical reasons, however, the individuals whose differences one from another have been measured often form groups

*That is, for a random sampling of them.

of a somewhat adventitious character. For example, the individual differences between one student and another in university classes in psychology have been measured, because such classes are at the psychologist's service. So also "college freshmen," or "children of a certain school grade," or "those who are willing to reply to a series of questions sent out by mail" are groups determined by convenience rather than by significance.

The group is often deliberately narrowed, as when the individuals are all insane, or all intellectually deficient, or all morally delinquent, or all men of science.

Furthermore, many of the facts concerning individual differences came as by-products of investigations of general mental laws or of group differences. For example, the early investigators of the least noticeable difference, of the time required for a simple reaction or for a reaction with discrimination, of the range of consciousness and the like, often considered the divergences of single men from the average of all men as "errors" due to chance conditions. These investigators took no interest in these individual divergences save as annoying hindrances to the exact formulation of constant laws of mental life. So also students primarily interested in the differences of man from woman have, without desire, got results concerning the differences of one man from another.

Finally the experiments made for us by the general conditions of life often provide data concerning the differences of individuals within a group constituted by some very complex circumstances. The differences of teachers in respect to salary, of college graduates in respect to general achievement in life, or of criminals in respect to the number and nature of their convictions by courts of law are samples.

As a result, there does not exist any study of the differences of a random sampling of all human beings in a single mental trait; or of a random sampling of individuals of the same sex; or of individuals of the same sex and age. There are, however, studies of the differences of a random sampling of individuals of the same sex and *approximately* the same age and race. The boys 12 years 0 months to 13 years 0 months old found in the schools of a German or English town would form such a group, and such boys have been measured with respect to certain mental traits. There are also studies of individuals of the same sex and of nearly the same age, race, and training (in certain particulars). The children just mentioned, if limited to a given grade or standard in school, would be thus nearly alike in respect to school training.

In spite of the difficulties of interpretation that arise from the mixture of unknown degrees of age, race and training, four general facts about individual variations in traits taken one at a time seem highly probable. (1) The variations are, in general, greater in acquired than in original traits. (2) They are, in general, greater in traits peculiar to man than in traits characteristic of all mammals. (3) The variations are usually, perhaps always, continuous. One grade or degree or amount is not separated from the next as ten men is separated from eleven men, but as ten pounds is separated from eleven pounds. (4) The variations usually cluster around one central tendency or "type."

As a preliminary to the discussion of these Laws of Mental Variations I present in Tables 44 and 45 and Fig. 52 samples of measurements of groups of individuals in traits taken singly.

From these tables one can calculate the amounts of difference existing and the relative frequency of each. For example, in Table 44 it appears that the two most unlike individuals of 37 chosen at random from women students of education differed by 656 seconds (896-240) in adding 48 examples, or by 13⅔ seconds per example. Differences of 12 seconds or over per example (576 seconds or over for 48 examples) occurred 11 times (between each of individuals a, b, c, and d and each of individuals J and K; also between each of individuals e, f and g and individual K); etc., etc. The median of all the individual differences will be found, by any patient reader who cares to compute it, to be about 3½ seconds per example. The median of the deviations of the 37 individuals each from 11 seconds

TABLE 44.

THE ABILITIES OF 37 ADULT WOMEN STUDENTS IN ADDING, AFTER ONE HOUR'S SPECIAL PRACTICE. TIME REQUIRED (IN SECONDS) TO ADD 48 SINGLE COLUMNS EACH OF 10 FIGURES TAKEN AT RANDOM FROM THE SERIES 2-9. 50 PER CENT OF THE TIME FOR ONE COLUMN WAS ADDED FOR EACH WRONG SUM.

Individual	Ability	Individual	Ability	Individual	Ability	Individual	Ability	Individual	Ability	Individual	Ability
a	240	h	390	o	465	v	508	C	669	J	845
b	242	i	428	p	482	w	535	D	719	K	896
c	249	j	433	q	488	x	545	E	729		
d	267	k	437	r	489	y	550	F	741		
e	272	l	459	s	499	z	572	G	763		
f	290	m	463	t	504	A	628	H	772		
g	315	n	464	u	506	B	642	I	811		

TABLE 45.

THE ABILITIES OF 37 ADULT WOMEN STUDENTS IN DRAWING LINES TO EQUAL 100 MM. LINES: AVERAGE ERROR (IN MM.) FROM THE STANDARD.

Individual	Ability	Individual	Ability	Individual	Ability	Individual	Ability	Individual	Ability	Individual	Ability
1	.9	8	2.0	15	2.8	22	3.4	29	4.9	36	9.1
2	1.3	9	2.2	16	2.9	23	3.7	30	5.2	37	10.2
3	1.5	10	2.3	17	2.9	24	3.8	31	5.5		
4	1.6	11	2.4	18	3.1	25	3.9	32	5.9		
5	1.8	12	2.6	19	3.1	26	4.2	33	6.0		
6	1.8	13	2.7	20	3.3	27	4.4	34	6.7		
7	1.9	14	2.7	21	3.4	28	4.6	35	7.3		

per example, which is the central tendency of adult women students of education in this trait, is 2.2 seconds per example.

THE AMOUNTS OF DIFFERENCE IN DIFFERENT TRAITS

Any such measurements of the frequencies of different degrees of difference between one individual and another or between individuals and some central tendency or type, are not, however, readily commensurate except within the same trait. To the question, "Do these women differ more in ability to add than in ability to accurately equal a length of 100 mm?," the answer is:—The range of difference is 13⅔ seconds in one case (5 to 18⅔), 9.1 millimeters in the other; the median difference is 3.5 seconds (approx.) in one case and 1.8 millimeters (approx.) in the other; the median deviation from the central tendency is 2.2 seconds in one case and 1.2 millimeters in the other.

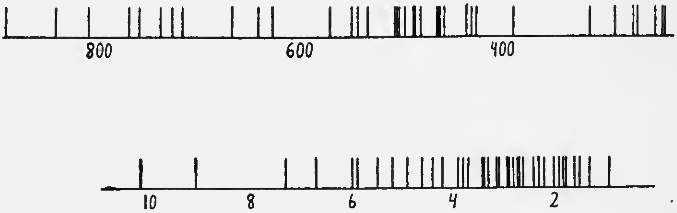


FIG. 52. (UPPER DIAGRAM). The Ability of Each of 37 Women in Adding. The scale gives the time in seconds required to add 48 single columns, each of 10 figures, 50 per cent of the time for one column being added for each wrong sum. Each vertical line represents one individual.

FIG. 52. (LOWER DIAGRAM). The Ability of Each of 37 Women in Drawing Lines to Equal a 100 mm. line. The scale gives the average deviation in millimeters from the standard. Each vertical line represents one individual.

Now in common speech, and in books on psychology and education as well, the differences between two individuals in two or more mental traits *are* compared. Such statements as, "John differs little from James in memory, and much in judgment," or "Men are much more alike in sense powers than in imagery," are made. If they are justifiable, the differences in different traits must somehow have been made commensurate.

In particular, my statement concerning the variations in original *versus* acquired traits, or in mammalian *versus* distinctively human traits, must depend upon some means of comparing the magnitudes of variations in different traits.

In comparing the tallest and shortest of 100 adult men, we can say, not only that they differ by, say, 20 inches, but also (supposing them to be 76 and 56 inches) that the tallest is one and five-fourteenths times as tall. In comparing the wealth of the men we can say not only that one has $10,000 more, but also that (supposing them to possess $1,000 and $11,000) he is eleven times as rich. In a certain real and useful, though limited, sense, it can be said that adult men differ more in wealth than they do in weight, and more in weight than they do in stature. The ratio of the difference of one man from absolute zero, or just not any amount of the thing in question, to the difference of another man from the same zero or just-notness is in a sense commensurate with a similar ratio in the case of another thing. Nothingness is taken to be the standard and 'just not any' of one thing is treated as equivalent to 'just not any' of another thing.

If one wished to use the test in addition to place a person on a scale for "ability in addition" one would find no such clear and sure way to make the judgment of "times as far from 0 ability." It would be very hard to define "ability to add which is just barely not no ability to add" in terms of any test's score. It would be still harder, if not impossible, to define it in terms of a score in this test. Should it be 100 sec. per example, or 1000, or 10,000, or 100,000, or infinity?

If one wished to use the test in drawing lines to place a person on a scale for ability to notice differences he would have not quite so hard a task. For there would be some reason for taking, as 0 ability, a divergence of 100 mm.,—that is, failure to distinguish from the standard a line of zero length.

There is great absurdity in common opinions of what "just not 0" is in the case of a mental trait and of what the relative differences between this and that manifestation of the trait and 0 are. Even gifted and well trained thinkers will assert, some that the best handwriting is only one and a half times as "good" as a nearly illegible scrawl, and some that it is eight or ten times as good. They will declare, some that the average man knows twice as much as the average dog, and some that he knows a thousand times as much. Some of them will assert, on finding that one boy spells 48 words correctly out of a list of 50 and another 6, that the former was thereby proved to be eight times as good a speller. Even if we should all study somewhat exhaustively the logic of the "times" judgment in mental traits, there would still be fairly wide disagreements concerning, say, how many times as much curiosity the most curious man had in comparison with the average man or in comparison with the least curious man; or concerning how many times as much mathematical knowledge the most learned mathematician had in comparison with the stupidest idiot.

But by any rational and just decision as to what 'o' or 'just not anything' was for each mental ability, we should find abundant evidence of the truth of the two laws—that individual differences in single traits are greater in acquired than in original abilities; and that they are greater in abilities peculiar to man than in abilities possessed by mammals in general.

For example, suppose that we regard as absolute zeros, the following: just not any A's marked in a minute, just not any sums correctly given in a minute, just not any dots placed in a series of squares, and the like, using as the measure in every case the amount done of some task, some amount of which nearly all adult Americans can do. The 'times as much' then comes out greater for thought than for movement, greater for memory of a passage than for memory of unrelated words or numbers, greater for responding to the meanings of words than for responding to the differences of colors, greater for marking words containing both a and t than for marking A's amongst other letters, greater in solving problems in arithmetic, mechanical difficulties, or geometrical puzzles than in sorting out colors.

For the reasons stated at the beginning of this chapter a survey of the data bearing on the comparative variability of man in different traits would not be desirable, even apart from the difficulty of choosing zero points. Moreover, the truth of the statement that as a rule individuals differ more in acquired than in original traits hardly needs statistical proof. The range of variations or the average individual difference is obviously greater in the acquired perceptions of words, music, geometrical forms and the like than in the original sensitivities to colors, sounds or distances; in the acquired movements of writing, sewing, singing and the like than in the original reflexes and instincts of winking, swallowing, clasping, running or striking; in knowledge, which is largely acquired, than in movement, which is to a considerable extent original; in the interests in literature, science or politics than in the interest in sex.

Within a narrow group, of course, uniformity of conditions may occasionally act to reduce natural differences, for example in habits of eating. If, however, men are taken over all the world and over a number of centuries it is hard to find any trait where the modifications by training do not increase natural differences.

The second law, that variations are greater in traits peculiar to man than in traits common to man and the mammals in general, is evidenced by the comparison of variability in remembering ideas with variability in remembering acts of skill, by the comparison of variability in marking A's with variability in marking logically absurd sentences, by the comparison of variability in drawing a line between two lines accurately with that in defining a word accurately, and the like. As the superiority of the best to the worst philosopher is greater than that of the best to the worst rememberer of places or avoider of animal enemies, so the variability in thinking with ideas in general is greater than that in the simple sensori-motor processes.

THE CONTINUITY OF MENTAL VARIATIONS

Continuity of variations means two things,—the absence of regularly recurring gaps, such as those between 2 petals, 3 petals, 4 petals, and the like, and the absence of irregularly recurring gaps, such as those between mice and rats, between rats and squirrels, and the like.

That continuity* of variations in a mental trait taken

*Of course continuity is not taken here in the sense of infinite divisibility. There are doubtless ultimately unit-factors which either act or do not act, singly is the rule can best be realized by trying to find exceptions to it. Such there may be, but I am not aware that any mental trait varying in amount has been shown to vary by discrete steps. A misleading appearance of regularly recurring gaps often arises from inadequate measurements. In a test of memory, for example, 12 nonsense syllables being read, individuals may appear in the scores as 5, 6, 7, 8 and 9 without any 5.5's, 6.75's and the like. But if four such tests are made and the average is taken, there will be 5.5's and 6.75's.

A misleading appearance of irregular discontinuity often arises from the insufficient number of cases measured. If only a few individuals are measured in a trait or if the scale is a fine one, there will of course be divisions on the scale or amounts of the trait unrepresented in any individuals. Thus in Fig. 52 there are such gaps, which would all have been filled in, had the number of individuals been very large. The 37 individuals whose abilities were reported in Fig. 52 were, in fact, chosen at random from 200. If all the 200 records had been used in constructing Fig. 52, its gaps would have been largely filled in. For example, the 37 cases show, between the abilities 500 and 700, a distribution as in the upper surface of Fig. 53, whereas the 200 cases show, over the same extent of the scale, a distribution as in the lower surface of Fig. 53.

It should be unnecessary to warn the reader against the absurdity of deliberately changing continuous variations into a few groups by coarse scaling; next assuming that the central part of one of these coarse divisions really measures all the individuals therein; and finally imagining that, because the continuous series, varying from a to $a + b$, has been called, say, Poor, Medium, Good and Excellent, there are really gaps

and which consequently increase the amount of the trait by either zero or a certain amount. But the discrete steps are exceedingly small like the steps of increase of physical mass by atoms. Intelligence, rate of movement, memory, quickness of association, accuracy of discrimination, leadership of men and so on are continuous in the sense that mass, ampèrage, heat, human stature and anemia are.

within it! Unfortunately even gifted thinkers are guilty of this error.

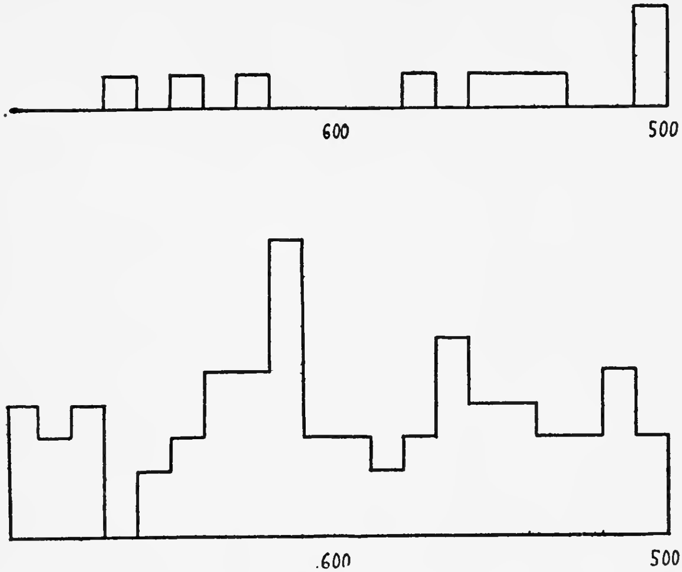


FIG. 53. The distribution of the cases falling between 500 and 700 seconds in adding 48 columns each of 10 one-place numbers, when, in all, 37 individuals were measured (upper diagram); and when, in all, 200 individuals were measured (lower diagram).

THE RELATIVE FREQUENCIES OF DIFFERENT AMOUNTS OF DIFFERENCE

The question of the clustering of variations around one central tendency demands more elaborate treatment. Fig. 54 shows the relative frequencies of the different amounts of the trait in the case of six mental traits. These six distributions illustrate the statement on page 317 that 'variations usually cluster around one central tendency.' This statement is not, however, universally, or even commonly, accepted. On the contrary the common opinion is that the distribution of individuals with respect to the amount of a single trait is multimodal, as in Fig. 55A, or even a compound of entirely distinct species, as in Fig. 55B. There would then be many small

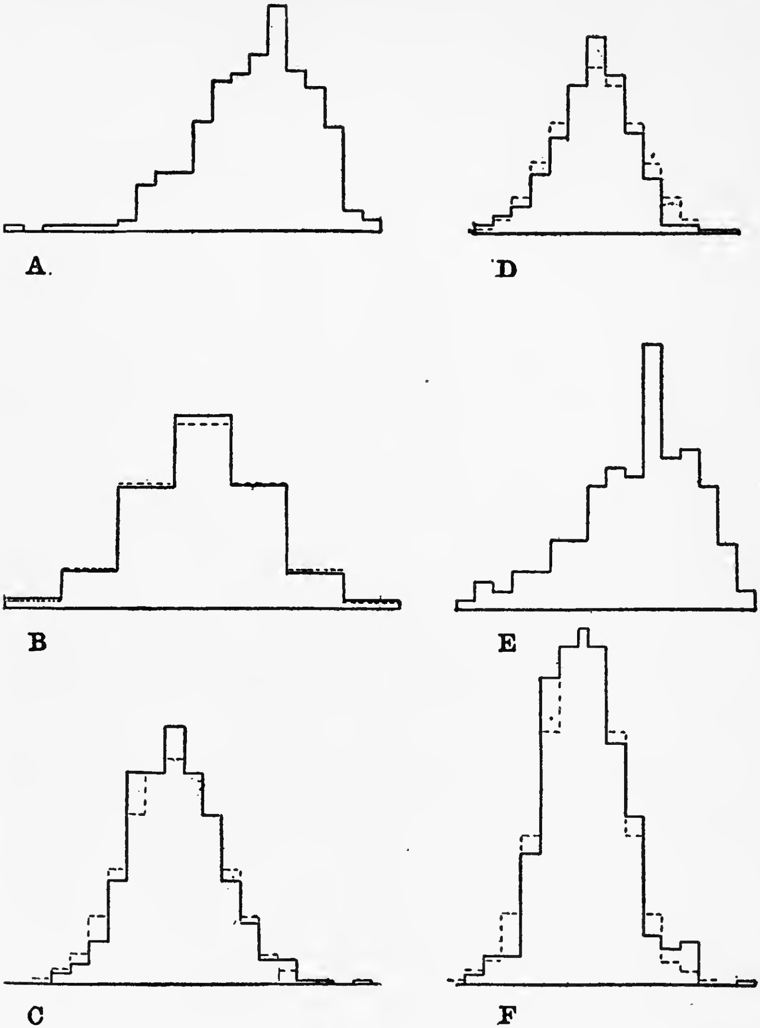


Fig. 54. Samples of the Forms of Distribution Found in Mental Traits.
A. Reaction time: 252 college freshmen.
B. Memory of digits: 123 women students.
C. Efficiency in marking A's on a sheet of printed capitals: 312 boys from 12 years 0 months to 13 years 0 months.
D. Efficiency in giving the opposites of words: 239 boys from 12 years 0 months to 13 years 0 months.
E. Accuracy in drawing lines to equal a 100 mm. line: 153 girls from 13 years 0 months to 16 years 0 months.
F. Efficiency in marking words containing each the two letters a and t: 312 boys from 12 years 0 months to 13 years 0 months.

In all six cases the left end of the scale represents the lowest abilities—that is, the longest times in A, the fewest digits in B, etc. The continuous lines give the distributions. The broken lines are to be disregarded for the present.

differences and many large differences with few cases of medium differences. This may be called the 'multiple type' theory. For instance, in the case of intellect we find the terms genius, normal, feeble-minded, imbecile and idiot used as if the

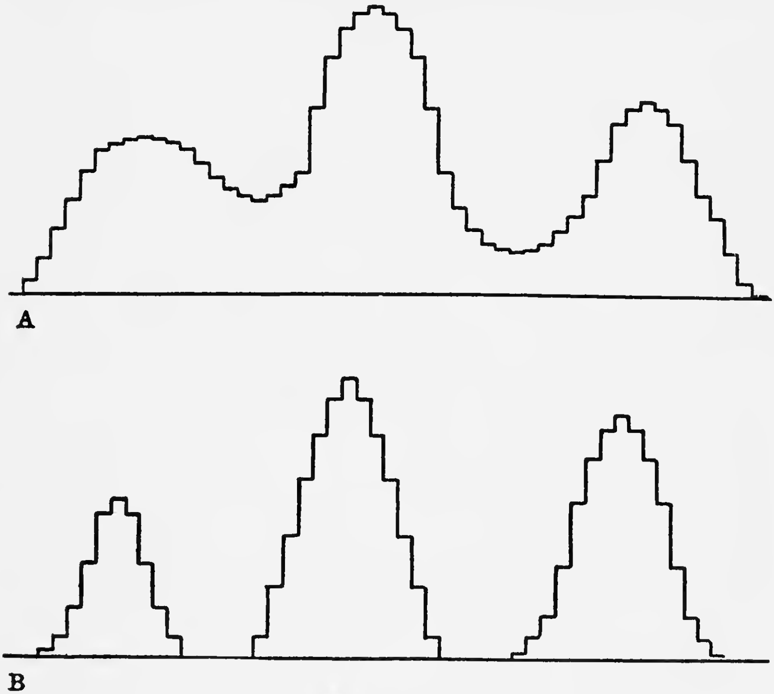


Fig. 55. Multimodal Distributions.

geniuses were separated by a clear gap from the normal individuals, these again from the feeble-minded, and so on. So also visualizers and non-visualizers, or men of normal color vision and the color blind, are spoken of as if those in each group were all almost identical and all much unlike all in the other group.

Multimodality is to be expected in traits the amount of which may be greatly increased or decreased by some one cause (or number of causes commonly acting together). If, for instance, reading Aristotle added enormously to anyone's intellectual gifts, we should expect to find men divided into two distinct surfaces of frequency on a scale for intellect, the higher ranking species being made up almost exclusively or even entirely of those who had taken the Aristotelian dose. The action of ophthalmia in causing blindness in the new-born is such a cause. By reason of it and other diseases, the visual capacity is reduced enormously in certain individuals, so that there are two modes, the seeing and the blind. Fewer men are able just barely to see than are totally blind. Injury to the head at birth or disease of the thyroid gland may be such a cause, reducing certain individuals all to an equal condition of intellect, so that possibly, by reason of it and other accidents and diseases, the number of very idiotic children may be greater than the number of those less idiotic over an equal length of the scale (see Fig. 56). This is doubtful, however.

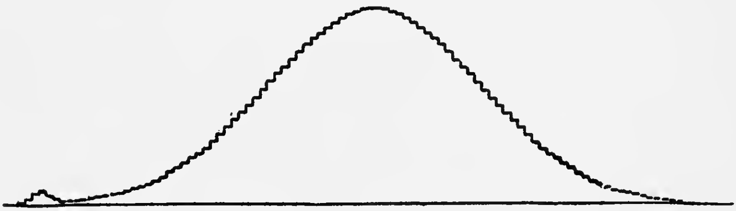


FIG. 56. A Distribution Showing a Secondary Type, of Great Inferiority.

In certain traits, such as knowledge of a certain language, or ability to play a certain game, there are two species. One includes those who have had no opportunity to get the knowledge or ability and whose knowledge or ability is consequently 0; the other is made up of those who have had some opportunity to get the knowledge or ability and who range in it from 0 or near 0 to a large amount. Understanding of spoken English, or ability to play chess or whist or golf, or ability to typewrite or to navigate a ship by the compass, would, of course, give such groups, if measured in adults the world over. Here the cause does not produce a uniform amount of the trait, but the world is so arranged that on many persons the cause does not act at all.

Many such causes may act in the case of particular habits, knowledges and skills. Since, for example, some Germans are, and some are not, subjected to the action of enforced military service, there may well be two modes to the surface of frequency of knowledge of the manual of arms, one group all

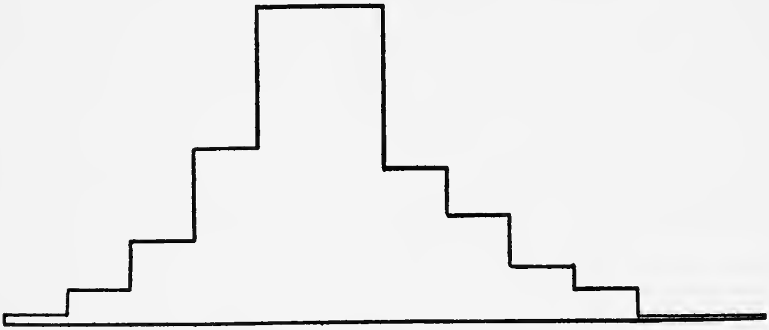


FIG. 57. Relative frequencies of different amounts of efficiency in marking A's on a printed sheet of capital letters, in the case of a group of twelve-year-olds, boys and girls together.

knowing it very well, the other group knowing hardly anything about it. Apprenticeship to a certain trade, or enrollment in a certain kind of school, may thus lead to extreme and uniform amounts of knowledge of, say, plastering or typewriting or

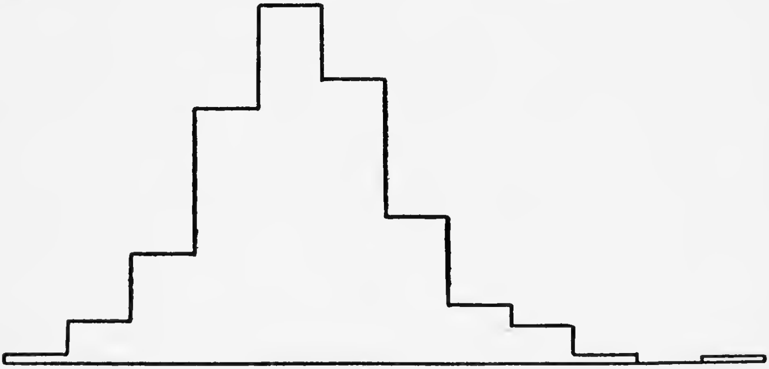


FIG. 58. Relative frequencies of different amounts of efficiency in marking A's on a printed sheet of capital letters, in the case of twelve-year-old boys.

medicine, so as to divide human nature sharply into an ordinary and an expert class. How far this happens is not known.

If sex made a great enough difference in the amount of any trait, there would be two modes in the surface of frequency for the trait in question in the two sexes combined. But observable bimodality as a result of mixture of the sexes does not in fact appear, because the sex differences are so small. For example, Fig. 57 shows the result of such mixture in one of the traits in which the sexes differ most. Figs. 58 and 59 show the distribution separately for each sex.

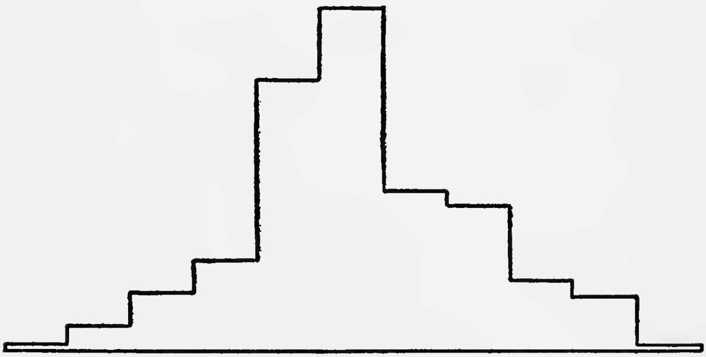


Fig. 59. Relative frequencies of different amounts of efficiency in marking A's on a sheet of printed capitals, in the case of twelve-year-old girls.

In traits in which race makes a great difference there will tend to be a mode for each racial type if two extreme races are mixed. Fig. 60 shows the results of so mixing civilized Europeans with Negritos (Figs. 61 and 62 giving the separate distributions for Europeans and Negritos). But if all races, or a random selection of races, were mixed, the resulting surface of frequency would not show a distinct mode for each, or probably for any one. Even so great a difference as that between the whites and the colored in scholarship in the high school is shown in the combined distribution only by a flattening of the surface of frequency as compared with that of either race alone (see Fig. 63).

In traits in which age makes a great difference, there will be a marked flattening of the surface of frequency. Thus, whereas for any one age, say 10 years and 7 months, the varia-

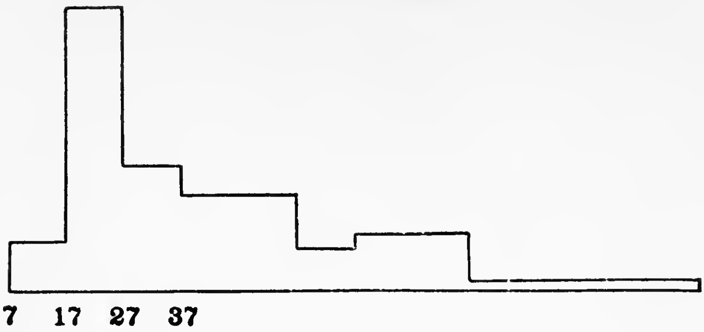


FIG. 60.

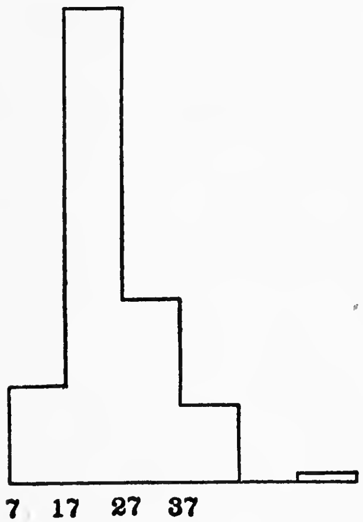


FIG. 61.

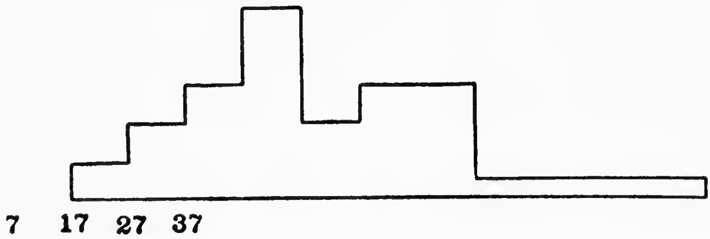


FIG. 62.

FIGS. 60, 61 and 62. Relative Frequencies of Different Times Required to Insert Variously Shaped Blocks in Holes to Match. The best abilities (that is, the shortest times) are at the left end of the scale in each case. FIG. 60 gives the distribution for a group equally divided between whites and Negritos. FIG. 61 is for whites alone. FIG. 62 is for Negritos alone.

tions with respect to 'School Grade Reached' will cluster closely around one grade, the distribution for individuals of all ages from 8 years through 14 years, shows three almost equally frequent degrees of the trait. This fact is shown in the case of the boys of one large city in Fig. 64.

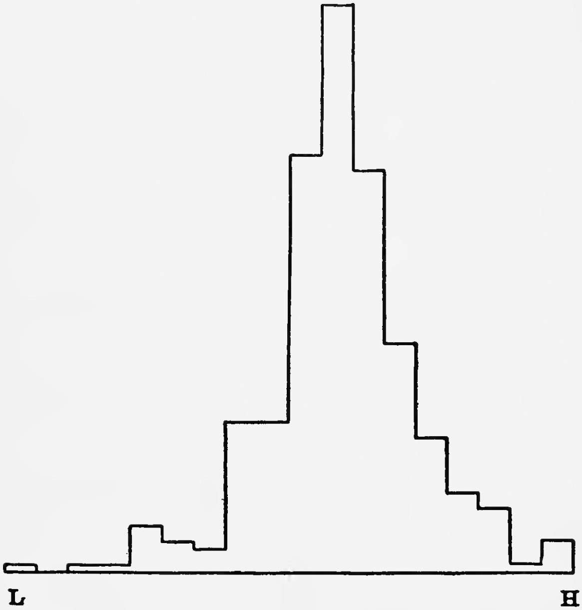


FIG. 63. The Relative Frequencies of Different Degrees of High-school Scholarship in a Group Composed of 150 Whites and 150 Negroes. The lowest grade of scholarship is at the left, the highest grade at the right, end of the scale. The two separate distributions, here combined, are shown in FIG. 43.

The common opinion that there are distinct species of individuals, with more or less pronounced gaps between, does not, however, limit itself to presupposing such multimodalities as those made by men and women, by Germans and Bushmen, by five-year-olds and fifteen-year-olds, by the ordinary population and the blind in respect to vision, by plumbers and non-plumbers in respect to skill in plumbing, by those who never tried to learn chess and those who did, in respect to ability at playing chess, and the like. It knows little or nothing of the effect of various combinations of causes upon the form of dis-

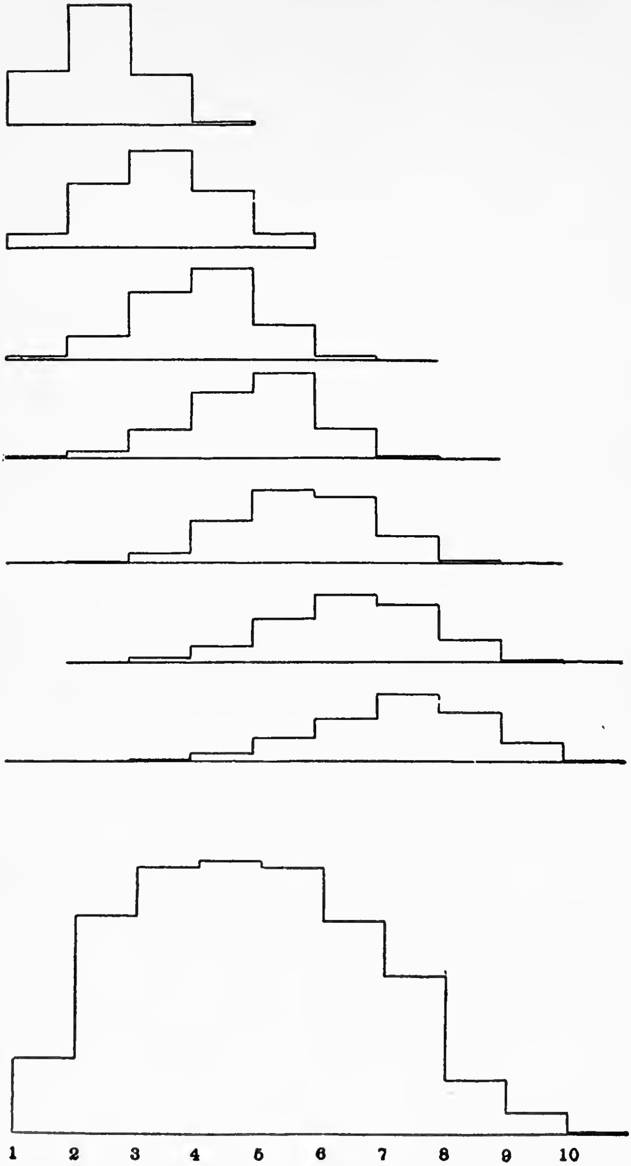


FIG. 64. The Relative Frequencies of Different Degrees of Progress in School in the Case of Boys 8 Years 0 Months to 14 Years 11 Months Inclusive. The horizontal scale at the bottom is for 'grade reached,' 1, 2, 3, etc., meaning first grade, second grade, third grade, etc. The surfaces of frequency are, in order from the top down, for 8-year-olds, 9-year-olds, 10-year-olds, 11-year-olds, 12-year-olds, 13-year-olds, 14-year-olds, and 8-14-year-olds combined.

tribution of a trait and it thinks of men as divided off into sharp classes in mental traits chiefly because it has not thought properly about the question at all. It merely accepts the crude adjectives and nouns which express primitive awareness of individual differences, as representatives of corresponding divisions in reality; neglects the existence of intervening

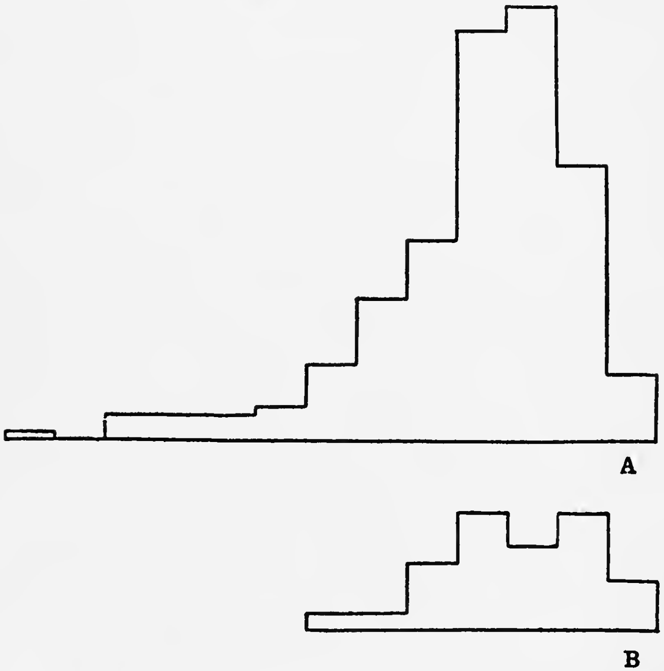
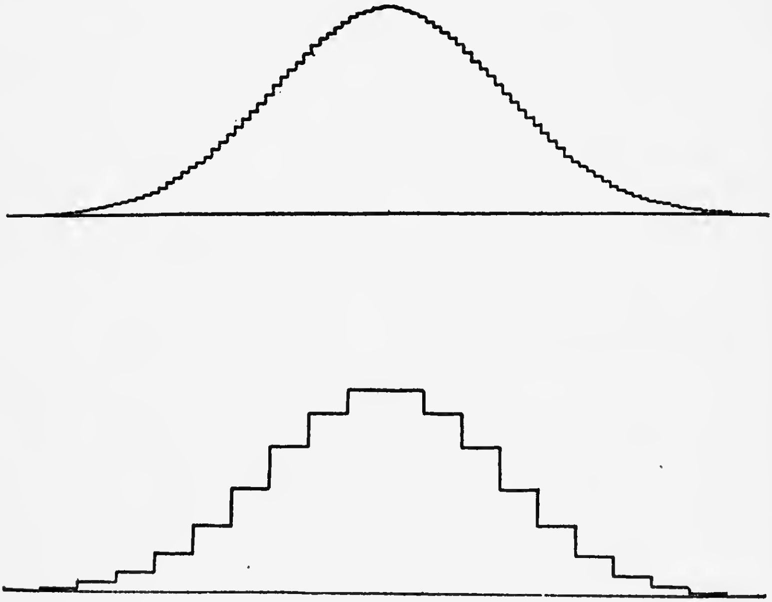


Fig. 65. The Distortion of the Form of Distribution Due to the Presence of Too Few Cases. A is the distribution as found from several hundred cases; B is that found in the same trait from the first 28 of them.

grades; and does not even attempt to estimate the frequencies grade by grade. How strong this tendency to verbal thinking is can be beautifully illustrated by the firm conviction of even long-trained men of science that people are either markedly right- or markedly left-handed, are either 'normal' in color vision or far removed from the 'normal' in color weakness or color blindness. Until recently the superstition that a great gulf separated children of normal intellect from the imbeciles and idiots was also very strong in many scientific men. The multiple-type theory does not refer to the separation of individuals into groups by the presence or absence of some one cause, or closely interrelated group of causes. It simply vaguely fancies that individuals, even of the same sex, race, age and training, somehow naturally fall into distinct classes or 'types.'



Figs. 66 and 67. The 'Chance' or 'Probability' Form of Distribution.

In such a form it is surely almost always, if not always, wrong. A group of such individuals does not, as a rule, show a separation into two or more groups, all in one being much like each other and little like any of those in the other group, or groups. Here again the rule may be verified by searching for exceptions to it. I know of no such. The misleading appearance of such may come either (1) from inadequate measurements, as in the case of the pseudo-discrete variations, or (2) from the examination of an inadequate number of individuals. Thus (1) if men are rated, for sensitiveness to red and green, as color blind, color weak and of 'normal' color vision, there will appear to be three types, just as there will appear to be three types of stature if all men are rated as short, medium or tall. But in careful tests the color blind will vary among themselves, the color weak likewise and the 'normal' likewise; the color blind will merge imperceptibly into the color weak and these into the 'normal.' Thus (2) if only fifty individuals are measured and if the scale is arranged so that these are included by, say, twenty divisions of it, there is a fair probability that, though the distribution of a thousand individuals show a clustering around one type as in Fig. 65A, the fifty may be clustered around two or more types as in Fig. 65B. The absurdity of inferring the existence of two species of human nature with respect to a trait from two apparent modes found when only a few individuals are studied should be obvious.

THE 'CHANCE' OR 'PROBABILITY' DISTRIBUTION IN THE CASE OF SINGLE MENTAL TRAITS

In sharp contrast to the common notion that human beings are divided sharply into classes, is the theory that, for any one trait, men, at least by original nature, form always only one class defined by one single sort of distribution. This theory answers in the affirmative the questions: "Do the distributions of mental traits in groups of individuals follow any regular law? Are the differences between individuals in mental capacities and characteristics amenable to any single type of description?" It supposes that, in all original traits, human beings so differ as to make the distribution that of a chance event, the surface of frequency being that of the probability integral. The exact meaning of this supposition and the basis for it need not now be discussed. Our present interest is in discovering how far any one type of distribution does characterize all original mental traits in human beings. By using graphic representations rather than algebraic formulæ, the answer and the evidence for it can be made clear even to one who knows nothing whatever of the mathematical properties of the surface of frequency of a chance event or of any other.

Fig. 66 gives the distribution or surface of frequency of the type to which by this theory the distributions of natural abilities conform. Fig. 67 gives the same distribution as Fig. 66, but with a coarser separation into grades.

Before comparing actual distributions with this theoretical form of distribution it will be profitable to inquire somewhat more systematically into the relation between the factors which

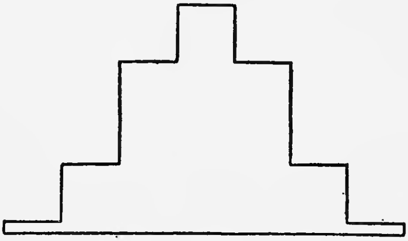


FIG. 68. The 'Chance' Form of Distribution from Six Equal and Independent Causes.

determine the amount of a trait in an individual and the relative frequencies of its different amounts among men.

Suppose the amount of a trait to be determined by six causes or factors—a, b, c, d, e, and f—each contributing 1 toward it; and suppose that each individual's nature includes a chance drawing from these causes, any combination being equally likely to be drawn.

The possible combinations are:—

None
a b c d e f
ab ac ad ae af bc bd be bf cd ce cf de df ef
abc abd abe abf acd ace acf ade adf aef bcd
bce bcf bde bdf bef cde cdf cef def
abcd abce abcf abde abdf abef acde acdf acef
adef bcde bcdf bcef bdef cdef
abcde abcdf abcef abdef acdef bcdef
abcdef

Consequently for every individual possessing 0 units of the trait there will be 6 possessing 6 units of it, 15 possessing 2 units, 20 possessing 3 units, 15 possessing 4 units, 6 possessing 5 units, and 1 possessing 6 units. The relative frequencies of the different amounts of the trait will be:—

Amounts of the trait.	Relative frequencies of these amounts.
0	1
1	6
2	15
3	20
4	15
5	6
6	1

Fig. 68 shows the form of distribution graphically.

Suppose that instead of six such causes there were twenty. The results of chance drawings from the possible combinations would then be to give, for every individual possessing 0 units of the trait, 20 possessing 1 unit, 190 possessing 2 units, 1140 possessing 3 units and so on, as listed below and shown graphically in Fig. 69.

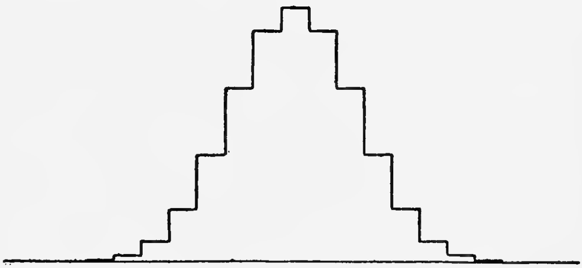


Fig. 69. The 'Chance' Form of Distribution from Twenty Equal and Independent Causes.

Suppose the number of causes to be increased, the other conditions remaining as before. The relative frequencies would assume more and more closely the proportions shown in Fig. 66 which is the surface of frequency of an event caused by a chance selection from among the combinations of a large number of factors each of small and all of equal amount. This surface is bounded by the probability curve. It is often called, or miscalled, the 'Normal' surface of frequency or 'Normal' distribution.

FORM OF DISTRIBUTION RESULTING FROM RANDOM COMBINATIONS OF 20 CAUSES OF EQUAL MAGNITUDE

Amounts of the trait.	Relative frequencies of these amounts.
0	1
1	20
2	190
3	1,140
4	4,845
5	15,504
6	38,760
7	77,520
8	125,970
9	167,960
10	184,756
11	167,960
12	125,970
13	77,520
14	38,760
15	15,504
16	4,845
17	1,140
18	190
19	20
20	1

The distribution of individuals in certain anatomical traits is much like that of Fig. 66. The close fit of individual variations in stature and the like to the particular bell-shaped surface of the probability integral naturally led to the expectation, or at least the hope, that all variable facts in original human nature would vary in this one way. In particular, individuals of the same sex, age, race and training should form a true species and vary in this one way.

As a matter of fact they do not, but they approximate it. No one form of distribution fits them all, but the bell-shaped curve given by the equation $y=e^{-x^2}$ fits them better than any other simple curve. FIG. 54 (p. 325) gives fair samples of the closeness of fit in mental traits.* The condition of an individual in a trait is very often the result of something approximating a chance selection of one out of many combinations of many factors, each of small and nearly equal influence on the trait's amount. In so far as it is so, the relative frequencies of the different conditions of individuals in respect to the trait will be approximately those of the so-called 'Normal' surface of frequency. But just as there cannot be two or more sharply defined types unless there are certain large causes whose presence makes a great difference in the trait's amount, so there cannot be a very close approximation to the 'Normal' sort of distribution unless the causes are nearly independent, and of nearly equal influence.

Such a trait as stature shows approximately the probability distribution, because its amount in any individual is determined by many only slightly correlated causes, such as the thickness of the skull, height of various bones of the head, thickness of each of the vertebrae, length of each of several bones in the leg, and the like. Each of these causes is again determined by very many imperfectly correlated causes. But such a trait as 'wealth possessed' will not. For certain relatively very large causes are at work, such as the discovery of a mine, or the death of rich relatives; and there is a strong correlation between many of the causes, for instance, between success in a business, directorship in a bank, advance knowledge of important actions, and power to influence these actions.

*The broken lines in FIG. 54 give, in each case, the 'chance' form of distribution.

"The form of distribution is then purely a secondary result of a trait's causation. There is no typical form or true form. There is nothing arbitrary or mysterious about variability which makes the normal type of distribution a necessity, or any more rational than any other sort or even any more to be expected on *a priori* grounds. Nature does not abhor irregular distributions.

On *a priori* grounds, indeed, the probability curve distribution would be exactly shown in any actual trait only by chance. For only by chance would the necessary conditions as to causation be fulfilled. And, in point of fact, as the reader has constantly been told by the adjective 'approximate,' the exact probability curve distribution does not appear in the facts or give signs of being at the bottom of the facts of mental life. The common occurrence of distributions approaching it is due, not to any wonderful tendency of a group of co-operating causes to act so as to mimic the combinations of mathematical quantities equal and equally probable, but to the fact that many traits in human life are due to certain constant causes plus many occasional causes largely unrelated, small in amount in comparison with the constant causes and of the same order of magnitude among themselves." [Thorndike, '04, p. 69 f.]

Distributions approximating it do occur very commonly in mental traits of original nature. And one will probably never be far misled by supposing that, in respect to the amount of original endowments in any trait, individuals of the same sex, race and age are distributed approximately according to the probability surface. The evidence from measurements points toward such approximation. Moreover, what is known of the physical basis of intellect and character leads to the expectation that many somewhat nearly equal factors are at work to determine the amount of any instinct or capacity possessed by men.

The meaning of some of the cases where the distribution of mental traits does not approximate the chance distribution will become clear if we examine, first, some cases of the distribution of a trait in a group of individuals of two or more distinct degrees of maturity or of training, and, second, cases where some selective agency has been at work. Fig. 70 gives the distribution of ability in the A test in a group of children 8, 9,

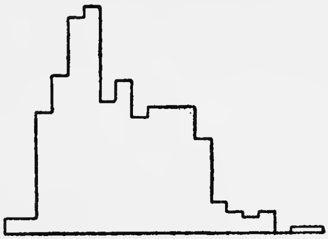


FIG. 70.

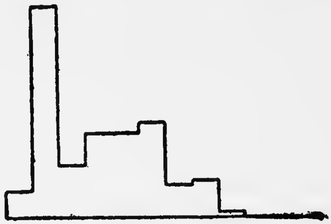


FIG. 72.

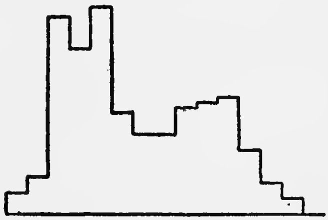


FIG. 71.

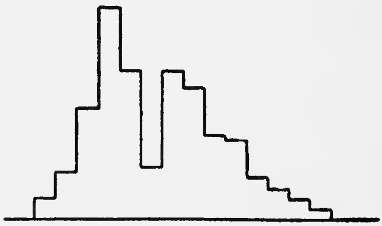


FIG. 73.

FIGS. 70, 71, 72 and 73. The Effect upon the Form of Distribution Due to Combining Groups Distinct in respect to Age, Training or Some Other Cause Affecting the Trait in Question.

FIG. 70. The distribution of ability, in the A test, of children 8, 9, 14 and 15 years old.

FIG. 71. The distribution of ability, in the opposites test, of children from grades 3 and 7.

FIG. 72. The distribution of ability, in multiplication, of children from grade 7 and from high school.

FIG. 73. The distribution of ability, in the a-t test, of children 9 and 15 years old.

14 and 15 years old. Fig. 71 gives the distribution of ability in writing the opposites of words in a group composed of about 140 girls in the third school grade and about 180 girls in the seventh school grade. Fig. 72 gives the distribution of ability in multiplication in a group composed of seventh grade pupils and high school pupils. Fig. 73 gives the distribution of ability in marking words, containing each the two letters a and t, in a group comprising 9-year-olds and 15-year-olds. In general when the surface of frequency of a mental trait departs from the probability curve toward a flattening and toward the

appearance of two or more modes, one may expect to find a mixture of sexes, races, ages or trainings.

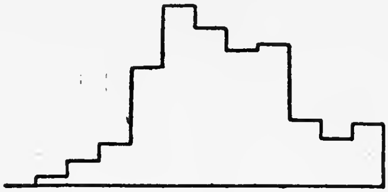


Fig. 74.

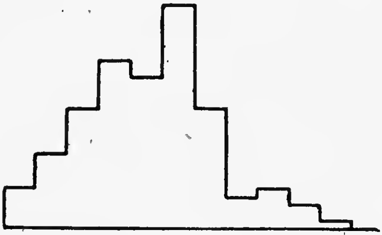


Fig. 75.

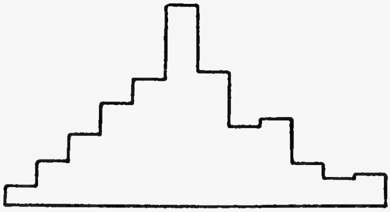


Fig. 76.

FIG. 74. The distribution of ability, in controlled association, of 12-year-old boys in the 6A grade or higher.
FIG. 75. The distribution of ability, in controlled association, of 12-year-old boys in the 6B grade or lower.
FIG. 76. The distribution of ability, in controlled association, of all 12-year-old boys in the school.
In FIGS. 74, 75 and 76, the left end of the scale is for the lowest degree of the trait.

Fig. 74 gives the distribution in a test of controlled association of the 12-year-old boys in the 6A grade or higher. The lack of symmetry in the surface is obviously due to the fact that we are dealing with a selected group—that the duller and less mature boys have been eliminated. The influence of the opposite sort of elimination is seen in Fig. 75, which gives the distribution, in the same trait, of 12-year-old boys in the grades lower than the 6A. By combining the two we have Fig. 76, which approximates the chance distribution. Figs. 77 to 80 give real cases of distributions distorted by selection. Fig. 77 shows the distribution of ability in giving the opposites of words in the case of twenty inferior day-laborers. Fig. 78 represents the ability of this same group of unintellectual and uneducated men in supplying words to fill omissions in a passage so that the total will make sense. Fig. 79 shows the dis-

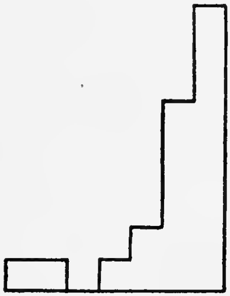


Fig. 77.

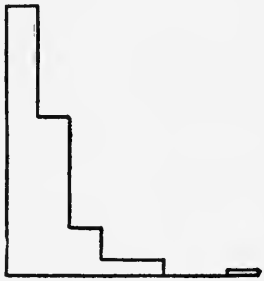


Fig. 79.

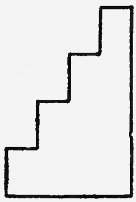


Fig. 78.

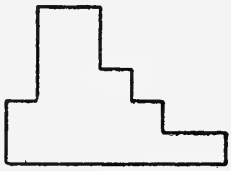


Fig. 80.

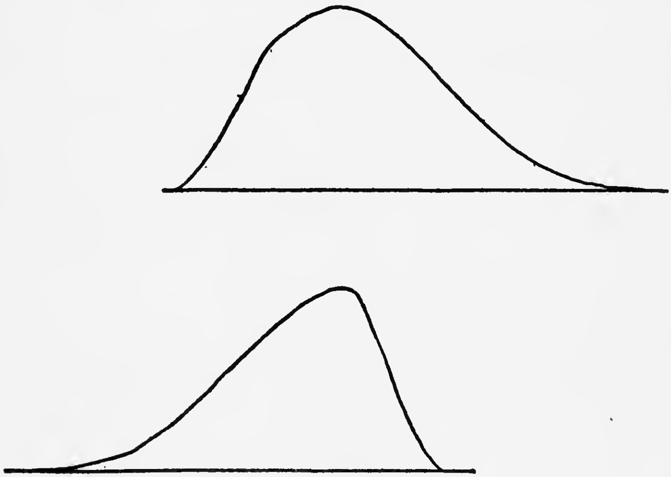
FIG. 77 shows the distribution of ability in giving the opposites of words, in the case of twenty inferior day-laborers.
FIG. 78 represents the ability of this same group of unintellectual and uneducated men in supplying words to fill omissions in a passage so that the total will make sense.
FIG. 79 shows the distribution of ability, in mathematics, of candidates for honors in Mathematics at Cambridge University.
FIG. 80 represents the distribution of ability, in seventeen university teachers and students, in checking the A's on sheets of printed capitals.
In FIGS. 77, 78, 79 and 80, the left end of the scale is for the lowest degree of the trait.

tribution of mathematical ability in the candidates for honors in mathematics at Cambridge University.* Of course such candidacy implies that the poorer grades of mathematical ability

*It is taken from Galton's *Hereditary Genius*, 2d ed. p. 16.

are eliminated. Fig. 80 represents the distribution of ability, in seventeen university teachers and students, in checking the A's on sheets of printed capitals.

Any selective agency which works upon a species of individuals will alter the shape of the surface of frequency for any mental trait, unless its selections are random with respect to different amounts of that trait. As the selective action is commonly such as picks out the good or the bad, the result is



FIGS. 81 (above) and 82 (below).

commonly to produce a "skewness" of the surface toward one extreme and a blunted condition at the other. When a series of measurements in a group of the same sex, age and training shows a deviation from the probability surface toward conditions like those of Figs. 81 and 82, it is wise to ascertain whether some selective agency has not been at work upon the group.

We have seen that the form of distribution of a single mental trait is usually such that the individuals cluster around one central tendency. We have seen also that, within a group of the same sex, race and age, in original traits, the variations from the central tendency occur in approximately the relative frequencies described by the probability distribution. A deviation of any degree plus is about as common as that of the same degree minus; or, more briefly, the distribution is approximately symmetrical. The average, median and mode therefore nearly coincide. The frequency decreases with the amount of deviation from the central tendency, at first slowly, then rapidly and then slowly. The average deviation from the central tendency is about 1.18 times the median deviation. About 82 per cent of the individuals differ from the C. T. by less than 2 times the median deviation; about 96 per cent by less than 3 times the median deviation; about 99½ per cent by less than 4 times it.

The most important fact, however, is not the commonness of this or that form of distribution, but the absolute law that the form of distribution is a result of the nature of the factors at work to produce the trait's amount. One large factor that is present for some individuals and not for others will always act toward the production of bimodality or distinct types. A multitude of nearly equal factors from which each man's nature and training is approximately a random selection will always act toward the production of unimodality, symmetry, and slow-rapid-slow decrease in frequency around the central tendency. An interdependence whereby the action of certain factors increasing the amount of the trait is dependent upon the action of other factors which of themselves bring the individual to a certain fairly high station, will always act to disturb symmetry. For every peculiarity in the causes determining a trait's amount in an individual there will be an effect in the relative frequencies of the trait's amounts in a group of individuals.

I should perhaps apologize to the reader for this long discussion of a matter which may seem to be sufficiently obvious without any discussion. I am glad if it does now seem obvious. The excuse for the long discussion is that the usages of language have persistently misled thinkers about human nature into supposing that it was, in each trait, divided into sharply separated classes to fit the adjectives and nouns by which primitive man roughly denoted different sections of a continuous scale, and that, on the other hand, certain thinkers have carelessly extended the particular 'probability' form of distribution to cases where it cannot possibly fit the facts.

CHAPTER XV

THE RELATIONS BETWEEN THE AMOUNTS OF DIFFERENT TRAITS IN THE SAME INDIVIDUAL

One feels a bareness and paltriness in such piecemeal descriptions of human beings and their differences one from another as have been given in the last chapter. The actual varieties of human nature do not stand out when one trait at a time is measured. Why, it may be asked, does psychology not take actual whole natures and state how they differ? Why does psychology not describe human minds as zoology describes animal bodies, by classifying them into families, genera and species, and by stating the differences between the different sorts of minds found?

It is true that zoology does not measure all animals in length, then in weight, then in color, then in number of organs, then in number of bones, and so on through a list of particular traits. It began with types or sorts apparent to common observation, such as worms and fishes, and described their essential features and the characteristic differences of one sort from another. And it is true that psychology might try to do likewise. If there were types or sorts of minds equally apparent to common observation, it would surely be worth while to start a description of human nature's varieties with them. But there are no sorts or types of minds that stand out clearly as birds, fishes and worms do amongst animal forms.* Psychology has first to find which the sorts or types are.

There are two ways of discovering them. The first is by

*The men allied by common ancestry, men of the same race, would be most likely to form a mental sort or type differing from other men, if not as fishes differ from other animals, at least as much as salmon differ from direct measurement of individuals *in toto,* and of their differences. To measure the difference between one whole man and another means to assign to each man his amount of each trait, and to measure each difference. The difference between the two men means just all those particular differences. To this method we shall return in the next chapter, though it may be said at once that no adequate measurements exist of even a single individual.

The second way of discovering the sorts or types into which men as total natures are divided is by discovering what each amount possessed of any one trait implies concerning the individual's condition in other traits.

A statement of the differences between one whole man and another would be an almost interminable inventory of particular differences, unless some traits were so related that knowledge of the amount of one of them possessed by a man informed us of the amount he possessed of the other also. Suppose, for instance, that any given amount of error in judging one length with the eyes always implied certain known amounts of error in judging all lengths, whether by eye, by arm-movement or by pressure, all weights, all colors, all pitches, all tastes, all smells, all brightnesses, all intensities of sound and all other sensory features of objects. In such case, one measurement would inform us, once for all, of a fairly large fraction of a man's nature and of his differences from another man, similarly measured. The necessary preliminary to the direct study of differences of total natures is thus the study of the relations of single mental traits.

The Measurement of Relations between Mental Traits

It is necessary to be clear at the outset in respect to just what is meant by the relation, or, as it is commonly called, the

other fishes. But even between races there is no surety that such is the case and no excuse for avoiding the slow and laborious comparison of individuals from different races in one after another trait that seems significant.

correlation, between two mental traits. It means, of course, the relation of some amount of one trait (A) to some amount of another trait (B). It also means, for the present purpose, the relation between an amount of A characteristic of a given individual to an amount of B characteristic of that same individual. The amounts might be the amount of A *more than zero* and the amount of B *more than zero;* or they might be the amount of A more than, or less than, some assigned amount and the amount of B more than, or less than, some assigned amount. Let us call the first sort *relations of divergences from zero* and the second sort *relations of divergences from arbitrary standards,* or *arbitrary relations.*

For example, suppose 5 eight-year-old girls, I, II, III, IV and V, in tests of memory of German equivalents of English words, to show the following abilities:—

Individual	Trait A Median number of words remembered (after a given amount (K) of training) at the end of 2 minutes	Trait B Median number of words remembered (after a given amount (K) of training) at the end of 60 days.
I.	23	11
II.	21	11
III.	16	9
IV.	11	3
V.	9	5

In the first meaning, the relation, of the divergences from zero is, of course, expressed as $\frac{11}{23}$, $\frac{11}{21}$, $\frac{9}{16}$, $\frac{3}{11}$ and $\frac{5}{9}$ respectively for these children. The central tendency of the relation is, roughly, to remember half as much after two months as was remembered after two minutes.

Suppose, now, that the relation sought is that between divergence from the central tendency for eight-year-old girls in A and divergence from the central tendency for eight-year-old girls in B, and that this central tendency is 15 for A and 6 for B. The relations are then $\frac{+5}{+8}$, $\frac{+5}{+6}$, $\frac{+3}{+1}$, $\frac{-3}{-4}$ and $\frac{-1}{-6}$.

Suppose that the relation sought is that between divergence from the central tendency of all human beings in A and divergence from the central tendency of all human beings in B, and that these central tendencies are 24 for A and 10 for B. The relations are then $\frac{+1}{-1}$, $\frac{+1}{-3}$, $\frac{-1}{-8}$, $\frac{-7}{-13}$, and $\frac{-5}{-15}$. It is clear that any relation varies according to the standards from which the related divergences are measured.

The relations to be considered in this chapter are arbitrary relations. We shall always be concerned with the relation of an individual's divergence from the central tendency of some defined group in one trait to his divergence from the central tendency of that same group in another trait.

The first fact to notice about such mental relations is that the same relation varies greatly in different individuals. Such a case as that of the last illustration, in which the relations were, for five individuals, —1.00, —.33, +.125, +.54, and +.33, is not at all exceptional. The relations of traits differ in individuals as truly as do the amounts of a trait. A single sample will suffice. Fig. 83 shows the relation of (A) divergence from the central tendency of adult college women in drawing a line* to equal a 50 mm. line to (B) divergence from the central tendency of the same group in drawing a line* to equal a 100 mm. line. Each dot represents the relation in one individual by its location. Each dot is below the point (on the scale for A) which is the individual's measure in A and opposite the point (on the scale for B) which is the individual's measure in B.

It is at once seen that superiority of an individual to the central tendency of adult college women in A does in general imply superiority to their central tendency in B, but that there

*Under certain specified conditions.

is a wide range in the relation in individual cases. If the relation had been the same in all individuals, the dots would all be on one line, since all the individuals under any one point of the horizontal scale (that is, of the same ability in A) would all be opposite the same point of the vertical scale.

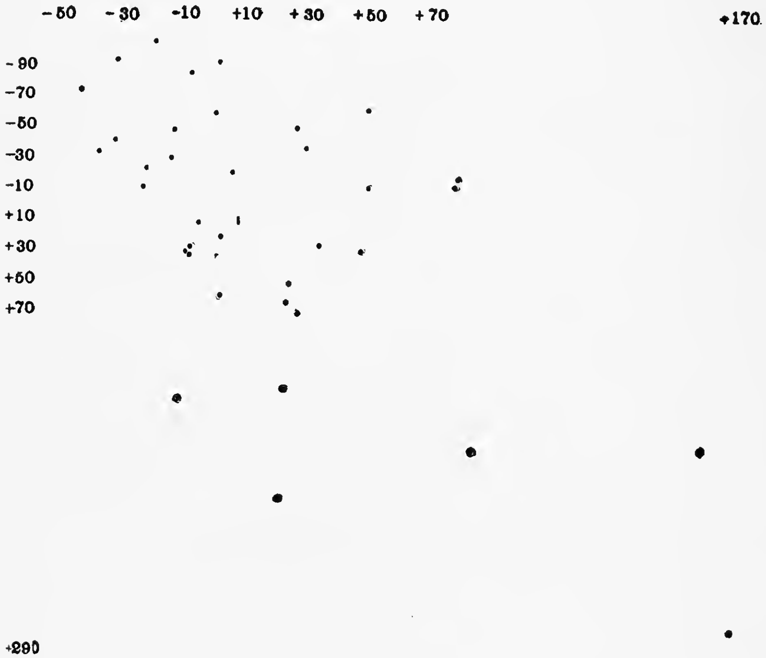


FIG. 83. The Relation between Ability in Equalling a 50 mm. Line (average of 30 trials) and Ability in Equalling a 100 mm. Line (average of 30 trials), in Women Students.

Since the relation does vary with individuals, it is fully measured or described only by such a list of all the individual relations as Fig. 83 gives. But its main features can be summarized, for any one degree of trait A, in two measures, one of the central tendency of all these individual relations, and the other of their variability around this central tendency.

Thus, suppose 10 individuals, all —8 in Trait A, to be, respectively, — 12, — 7, — 5, — 5, — 4, — 4, — 2, — 1, 0 and +6 in Trait B. The median of the ratios $\frac{-12}{-8}$, $\frac{-7}{-8}$, $\frac{-5}{-8}$, etc. is $\frac{-4}{-8}$, or + .5. The central tendency of the relation is for —8A to imply — 4B. The variability ranges from an implication of — 12 to one of + 6.

The relation may also vary with different individuals according to the amounts of A which they have. Thus, suppose

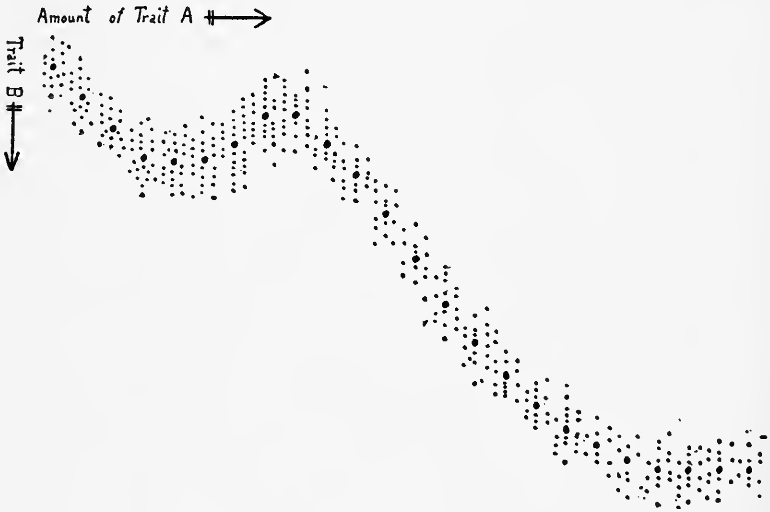


FIG. 84. A relation varying much with different amounts of Trait A, but varying little amongst individuals of like ability in A.

that the central tendency in the case of individuals all having +16A was to the relation $\frac{-2}{+16}$. The + .5 which restated the central tendency of the relation for individuals of — 8A, is replaced by — .125.

Fig. 84 shows a case (hypothetical) of a relation varying much with the amount of trait A, but with very little variation amongst individuals of like ability in A.

Fig. 85 shows a case (also hypothetical) of a relation varying little with the amount of trait A, but much amongst individuals of like ability in trait A.

Fig. 86 shows a case (also hypothetical) of a relation varying much in both respects.

In these figures the line formed by the points representing each the central tendency of the relation in the case of one amount of A may be called the *relation line*. In Figs. 84, 85 and 86, it is the line formed by the large dots.

When a relation varies according to the amount of one of the traits,—when, that is, the central tendency of the ratio

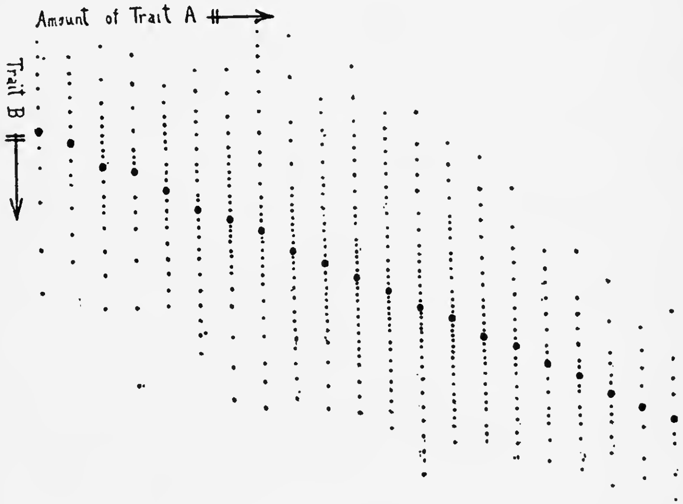


FIG. 85. A relation varying little with different amounts of A, but much amongst individuals of like ability in A.

A/B is different for A=k, A=2k, A=3k, etc.,—it may be called nonrectilinear, because the relation line is curved or broken. When a relation varies with individuals irrespective of their amounts of Trait A, the central tendency of the A/B ratio being the same for all values of A, it may be called *rectilinear* because the relation line will be a straight line.

When only a few individuals are studied, so that the number at any given point on the scale for A is not enough to give reliably the central tendency of the relation at that point, the relation will appear non-rectilinear even though it may really be rectilinear. For example, the relation between ability in the

A test and ability in the a-t test in 12-year-old boys (divergences from C. T.'s for that age and sex being related) is really rectilinear or very nearly so, but the relation line for any 100 boys will seem to be much broken as in Fig. 87. Whatever be the real form of the relation line, chance deviations from

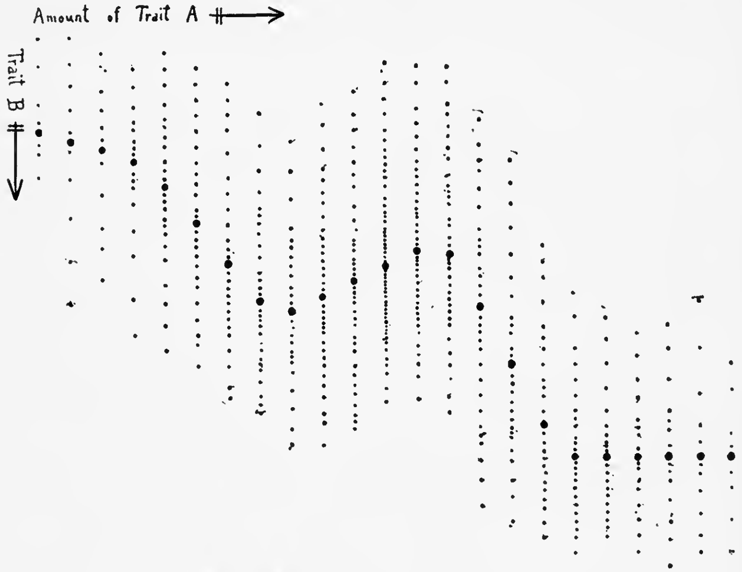


FIG. 86. A relation varying much in both respects.
FIGS. 84, 85 and 86. The variability of mental relations.
Trait A is scaled horizontally, the lowest degree being at the left.
Trait B is scaled vertically, the lowest degree being at the top.
Each dot represents the relation in one individual, the amount of A possessed by the individual being that represented by the point on the scale for A above the dot, and the amount of B possessed by the individual being that represented by the point on the scale for B at the left of the dot.
The large dots represent, each by its position, the central tendencies (medians) of those of each degree of ability in A.

that form will appear unless the number of cases is infinitely large. But in reality sharply irregular relations with zigzag relation lines,—such relations, for instance, as .5 for $A = k$, .7 for $A = k + 1$, .3 for $A = k + 2$, .5 for $A = k + 3$, .4 for $A = k + 4$, .6 for $A = k + 5$, .5 for $A = k + 6$, and so on—probably do not exist between any mental traits.

Relationships with curvilinear relation lines deviating markedly from straight lines,—such as that in Fig. 84 or those in Fig. 88,—may exist, but no such relation has as yet been proved to exist. It is therefore customary to treat relations between mental traits as approximately rectilinear.

The relation of two traits in a given group of individuals being expressed as the rectilinear relation which best fits them, their relation can be compared with that of any other traits in the same or any other group—provided one more simplification

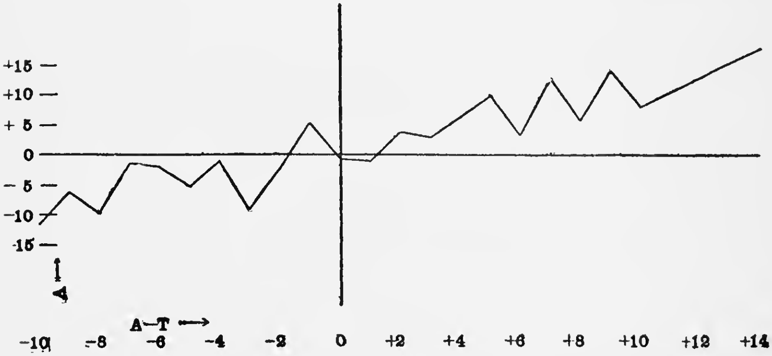


FIG. 87. The relation line in the case of the relation between ability in marking A's (one trial of 60 seconds) and ability in marking words containing both a and t (one trial of 120 seconds), in the case of girls of the same school grade. (With more adequate tests the relation would be much closer and the relation line somewhat straighter).

be made. This is that the divergences to be related be expressed each as a fraction of the variability of the trait in question. For example, suppose A, B, C and D to be series of divergences, in height (in centimeters), weight (in kilograms), number of words remembered in a certain test, and number of A's marked in the A test, from the condition of the modal twelve-year-old boy. Suppose that the central tendencies of the ratios B/A, C/A and D/A were respectively .5, .2 and .5. That would mean that a boy ten centimeters above the central tendency would in the long run be 5 kilograms heavier, remember 2 words more, and mark 5 A's more, than the modal twelve-year-old boy. This does not enable one to tell whether B/A is really equal to the relation D/A, though each is .5. Nor can one say with any useful meaning that B/A is two and one-half times as close a relation as C/A.

But if, instead of calculating the gross B/A, C/A and D/A ratios (.5, .2 and .5), we calculate the values of these ratios after every value of A is expressed as a multiple of the vari-

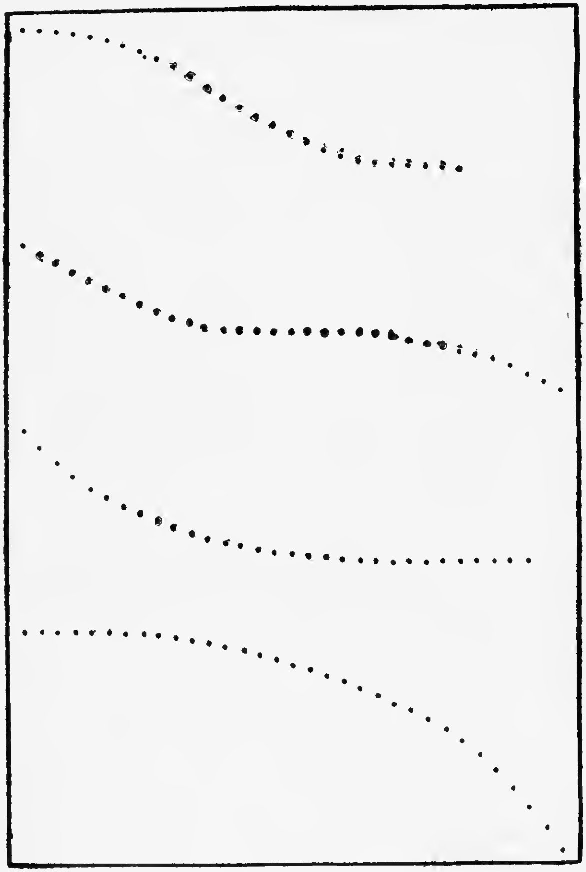


Fig. 88. Samples of curvilinear relation lines.

ability of twelve-year-old boys in trait A, every value of B is expressed as a multiple of the variability of twelve-year-old boys in trait B and so on, the ratios become commensurate.

Suppose the variabilities of twelve-year-old boys in the four traits to be, respectively:—Var. in A, 5 centimeters; var. in B, 4 kilograms; var. in C, 2 words; var. in D, 5 A's marked. Then the central tendencies of the ratios, $\frac{B/\text{var. } B}{A/\text{var. } A}$, $\frac{C/\text{var. } C}{A/\text{var. } A}$ and $\frac{D/\text{var. } D}{A/\text{var. } A}$ will be .625, .5 and .5.* Each of these central tendencies will measure the central tendency of the relation of (I) an individual's divergence in one trait measured as a multiple of the general tendency to diverge in that trait, to (II) his divergence in another trait, measured likewise as a multiple of the general tendency to diverge in that trait.

It is with such a meaning that we compare the closeness of the relation between, say, memory for numbers and memory for words with the closeness of the relation between memory for the same data over a short and over a long interval; or compare either of these with the relation between accuracy of discrimination of length and accuracy of discrimination of weight.

If each divergence is expressed in terms of the variability of its trait as a unit, we can think of an individual's condition in one trait as resembling or differing from his condition in another trait. The technique of measuring this resemblance between two traits in one man is then the same as that made familiar in Chapter XI in the case of the resemblance of two individuals in the same trait. *r*, the coefficient of resemblance, then measures the central tendency to resemblance or mutual implication found in pairs of amounts or conditions in two traits, each pair being characteristic of an individual. The amounts are divergences from the central tendencies of some defined group. The relation or resemblance or mutual implication is supposed to be constant for all amounts of either trait. *r* is a figure so calculated from the individual records as to give the one degree of relationship between the two traits which will best account for all the separate cases in the group. In other words it expresses the rectilinear relation from which the actual

*$\frac{B/\text{var. } B}{A/\text{var. } A} = \frac{B}{A} \times \frac{\text{var. } A}{\text{var. } B}$, or $.5 \times \frac{5}{4}$, or .625. Similarly for the $.2 \times \frac{5}{2}$ and for the $.5 \times \frac{5}{5}$.

cases might have arisen with least improbability. It has possible values from +1.00 through 0 to — 1.00. '$r = +1.00$' means that the individual who is the best in the group in one ability will be the best in the other, that the worst man in the one will be the worst in the other, that if the individuals are ranked in order of excellence in the second, the two rankings will be identical, that anyone's divergence in the one will be identical with his divergence in the other (both being reduced to multiples of the variabilities of the abilities to allow comparison). '$r = -1.00$' would, *per contra*, mean that the best person in the one ability would be the worst in the other, that any degree of superiority in the one would go with an equal degree of *inferiority* in the other. '$r = +.62$' would mean that (the two series of divergences being reduced to multiples of the variabilities in question) any given divergence in the one trait would imply, on the average, 62 hundredths as much divergence in the other.

THE RELATIONS BETWEEN THE AMOUNTS OF DIFFERENT TRAITS IN THE SAME INDIVIDUAL

As samples of *r*'s so calculated I give in Table 46 some of those obtained by Burt ['09] in the case of 30 boys between 12½ and 13½ years of age in an Elementary School in Oxford, attended by boys 'of the lower middle class,' whose parents paid 9d. a week tuition. The relations are all between divergences, each from the central tendency of the group in the trait in question. The traits were (in part) :—

Touch.—Delicacy of simultaneous discrimination of two points on the skin.

Weight.—Delicacy of discrimination of lifted weights.

Pitch.—Delicacy of discrimination of pitches (from 320 vibs.).

Length.—Delicacy of discrimination of length; deviation from the standard, 100 mm.

Tapping.—The number of holes made in 15 sec. in a sheet of paper over cloth, a needle fastened in a holder being used.

Dealing.—The time required to deal 50 cards into 5 heaps in the ordinary manner.

Sorting.—The time required to deal 50 cards into 5 heaps in accordance with the color of the cards. An error was not corrected by the boy, but was allowed for by an addition to his time-score.

Alphabet Sorting.—52 cards, each 20 mm. square and printed with a letter, each letter appearing twice in the series, were exposed in an irregular order. The boy was required to find a, then to find b, then c, etc., placing them in order in two rows. He then repeated this selecting again. The time taken was the score.

Immediate Memory.—Of 90 words (30 concrete, 30 abstract, and 30 nonsense) all of one syllable, each one seen and pronounced by the boy, given in series of 4, 5, 6, 7 or 8. For method of scoring see Burt, '09, p. 142.

TABLE 46.

THE INTERRELATIONS OF NINE MENTAL TRAITS IN THE CASE OF THIRTY BOYS. EACH ENTRY GIVES THE RELATION OF THE TRAIT LISTED ON THE LINE TO THE LEFT OF THE ENTRY TO THE TRAIT LISTED ABOVE THE ENTRY. ADAPTED FROM BURT ['09]. DECIMAL POINTS ARE OMITTED.

	Touch	Weight	Pitch	Length	Tapping	Dealing	Sorting	Alphabet	Memory
Touch		37	—01	26	12	32	20	66	15
Weight	37		06	-16	11	05	23	26	07
Pitch	—01	06		05	53	41	25	68	19
Length	26	—16	—05		16	59	32	37	05
Tapping	12	11	53	16		79	78	67	01
Dealing	32	05	41	59	79		77	83	18
Sorting	20	23	25	32	78	77		83	27
Alphabet	66	26	68	37	67	83	83		47
Memory	15	07	19	05	01	18	27	47	

The figures of Table 20 show the traits measured by the tapping, dealing and sorting tests to be so related that the amount of any one possessed by a boy gives a fairly close prophecy of the amounts that he will possess of the others. The traits measured by the tests of discrimination of touches, weights, pitches and lengths on the other hand are shown to be highly independent one of another.

Mental relations have been measured by Wissler ['01], Spearman ['04a, '04b, '06 (with Krueger)], Thorndike ['03, '09], Pearson ['07], Brown ['09, '10], Burt ['09], Bonser ['10], Frost,* Abelson ['11], Hart and Spearman ['12], Simpson ['12], and others. The results form a body of facts of great importance to psychology and education, but a summary of them without elaborate discussion of the methods used in each case, would be misleading. I shall, therefore, state only the general conclusions to which they lead.

The significance of the relations between mental traits which have been measured in this way is seen most easily and clearly by observing the doctrines about individual psychology which they disprove.

First may be mentioned a series of beliefs in mental antagonisms or compensations. Such are:—that superiority to the central tendency in vividness and fidelity of imagery of one sort implies inferiority to the central tendency in vividness and fidelity of imagery of other sorts; that superior ability to get impressions through one sense is related to inferiority in getting impressions through other senses; that intensity of attention varies amongst individuals in opposition to breadth of attention, so that a high degree of power to attend to one thing at a time goes with a low degree of power to attend to many things at once; that the quick learner is the poor rememberer; that the man of great artistic gifts, as in music, painting or literary creativeness, is weak in scientific ability or matter-of-fact wisdom; that divergence above the mode in power of abstract thought goes with divergence below the mode in thought about concrete things; that the man of superior intellect is likely to be

*In a study as yet unpublished.

of inferior mental health; that the rapid worker is inaccurate; that an agile mind goes with a clumsy body, etc., etc.

Not all of these and other supposed antagonisms or inverse relations have been specifically tested by the calculation of the appropriate *r*'s; but those which have been so tested have been found in gross error. Betts ['09] found the *r*'s for any one sort of non-verbal imagery with any other, in respect to vividness, completeness and detail, to be not only positive but high. Pedersen's data ['05] show, though he apparently did not notice the fact, a very close correlation between ability to grasp presentations through the eye and ability to grasp presentations through the ear. Meumann ['07] somewhat grudgingly admits that his contrast between "men with typically concentrative or intensive attention" and "individuals with typically distributive attention" ['07, vol. I, p. 500], does *not* mean that the person who is superior in the one is inferior in the other; that, on the contrary, he is more likely to be superior in the other also ("je grösser die Konzentrationsfähigkeit eines Menschen, desto grösser seine Distributionsfähigkeit" ['07, vol. I, p. 502 f.]) The author ['08] has shown that the individuals who learn a thousand words, as in a vocabulary, more quickly than the modal man, also, in a majority of cases, remember more of them after 40 days. Cattell ['03] finds that eminence in artistic lines implies superiority in politics or generalship or science more often than the reverse. All relevant measurements witness to a positive correlation between efficiency in thought with abstract data and efficiency in thought with concrete data; also between the ability to work with greater speed at a given accuracy, and the ability to work with greater accuracy at a given speed. Indeed the individual who works at higher speed often works more accurately at even that higher rate than does the slower worker at his more favorable rate.

The relations which do seem to be inverse are very instructive. They are mostly cases of the relation of a desirable divergence in one trait to an undesirable divergence in the other. Thus general intellect seems to be antagonistic to sullenness. Intellectual efficiency seems to be antagonistic to emotionality in the crude sense.

It is very, very hard to find any case of a negative correlation between desirable mental functions. Divergence toward what we vaguely call better adaptation to the world in any respect seems to be positively related to better adaptation in all or nearly all respects. And this seems specially true of the relations between original capacities.

The negative values in Burt's results, listed above, are probably not exceptions to this rule, but are due to the chance variations from truth which are to be expected from a small series of measures. His —.16 for the relation of discrimination of length to discrimination of weight, for instance, should be considered in connection with the +.52 obtained by the author ['09] for the same relation in the case of 37 young women, and the +.25 obtained in the case of 25 high school boys.

There may, however, be cases where some one large environmental agency acts to bring all those individuals subject to it up in one trait but down in another. The Roman Catholic Church might thus, at least at certain periods of its history, make many of its members more interested in theological argumentation, and less interested in scientific verification, than the modal man. The theory of asceticism might make all its adherents successful in contemplation and inefficient in action. City life might so stimulate adroitness in dealing with people and inhibit adroitness in dealing with animals and plants, as to produce a negative relation. Wherever some potent circumstances act to elevate one trait and depress another in individuals subjected to them, whereas alternative circumstances act in the opposite direction, there is a chance for individuals to show negative *r*'s in the traits in question, even though, in original nature, the traits are related positively. Such negative relations might appear, if men were measured from different national cultures or over thousands of years so as to include such contrasting environments, in cases where our present measurements show even strong positive relations.

Little is known about the shares of original mental organization and environmental influences in producing the relations of divergence in one trait to divergence in others, and still less about the special effect of what may be called irrational environments, those which weaken or inhibit certain generally desirable mental traits.

Finally, there are the notable cases of apparent compensation due to the special practice of one trait to make up for irremediable weakness in some other. Keenness of touch in the blind, is, of course, the clearest case. Of such cases it may be said that they are, for the general relation, trifles, and are partly balanced by other cases where special defect and special practice work to the opposite effect upon the total relation. It should also be noted that in original nature the rule is correlation, not compensation. Those defective in vision are, by nature alone, more likely to be defective in other senses than are those superior in vision.

On the whole, negative correlations between different 'efficiencies' or 'adaptabilities' or 'desirable traits' are surely rare, seem almost never to occur as a result of original mental nature, and require as causes peculiar oppositions of influence upon the two traits from the environment.

A second error in opinions about mental relations is in sharp contrast to the one just described. It is the doctrine that some one function is shared by all intellectual traits, and that whatever resemblances or positive correlations the traits show are due to the presence in each of them of this function as a common factor. In so far as they have it, they are identical. In so far as they lack it, they are totally disparate. In the words of Spearman ['04 b, p. 84], "All branches of intellectual activity have in common one fundamental function (or group of functions) whereas the remaining or specific elements of the activity seem in every case to be wholly different from that in all the others."*

*This quotation is unfair to Spearman, at least to his present view, but is used because it so clearly states an important extreme view. Writing

This doctrine requires not only that all branches of intellectual activity be positively correlated, which is substantially true, but also that they be bound to each other in all cases by one common factor, which is false. The latter would require that no two intellectual abilities or branches of intellectual activity should be more closely related to each other than to the fundamental function by which alone they are supposed to

with Hart in 1912, Spearman says: "Here, however, consideration must be turned to a point which appears to have given rise to misunderstanding. The opponents of the theory of a General Factor have taken this as claiming to be the *sole* source of correlation. Such an absurd claim does not seem really to have been advanced by any one. The earliest anouncement of the principle was accompanied by a warning of 'its inevitable eventual corrections and limitations.' Special emphasis was laid on the fact that correlation between performances is also produced by great similarity between them. Obviously, as the similarity tends towards completeness, the correlation must tend towards unity. This fact was underlined by actual examples in numerical detail. For instance, the correlation between Latin translation and Latin grammar was shown to be far too large to fit into the theory, and this was attributed to the content being the same in both cases, namely, Latin. Another instance of the same sort was French prose and French dictation. A further one was furnished by the test of counting letters one at a time and that of counting them three at a time; here, there was a close similarity both of content and form, and accordingly this was pointed out as the cause of the principle becoming invalid.

"It was never asserted, then, that the General Factor prevails exclusively in the case of performances too alike: it was only said that *when this likeness is diminished (or when the resembling performances are pooled together), a point is soon reached where the correlations are still of considerable magnitude, but now indicate no common factor except the General one.* The latter, it was urged, produces the basal correlation, while the similarities merely superpose something more or less adventitious. Up to the present, however, these similarities seem to have exercised surprisingly small influence. In all the performances dealt with in the next section, only three times did any of them resemble each other closely enough to require pooling; these cases will be discussed in detail later on." [Hart and Spearman, '12, p. 58 f.]. The view, so stated, does not differ from the view to be presented in this volume, save in the intimation in the last sentence that the 'similarities' other than in the general factor's strength, are of trivial effect in comparison with it. Later he says ['12, p. 76]: "The present paper, then, is far from denying that specific abilities may play a large role."

be related; and, as a corollary of this, that no four such abilities, A, B, C and D, should be more closely related in pairs, A to B, and C to D, than the element common to A and B is related to the element common to C and D. But unless one arbitrarily limits the meaning of 'all branches of intellectual activity' so as to exclude a majority of those so far tested, one finds traits closely related to each other but with their common element only loosely related to the common element of some other pair.

The next error may be roughly described as the supposition that for any one operation that is the same in *form*, such as discrimination of differences, attention, observation, inference or the like, the varieties produced by different *data* or *content* are perfectly correlated.

It has been common in psychological and educational literature to presuppose that the functions which we group under the same name, *e. g.*, attentiveness, somehow implied each other, that, for instance, a high status in attentiveness to school work was closely related to a high status in attentiveness to social duties, business pursuits, mechanical appliances and all the other facts of the individual's experience. Our rough and ready descriptive words, such as accuracy, thoroughness, reasoning power and concentration, have been used as if the quality must be present in approximately equal amounts in all the different spheres of mental activity. The notion of any special mental act, *e. g.*, the discrimination of 100 millimeters from 104, has apparently been that some general faculty or function, discrimination, was the main component, the special circumstances of that particular act being very minor accessories. Thus all the different acts in the case of discrimination would be very closely related through the presence in them all of this same mental component.

On the contrary, measurements reveal a high degree of independence of different mental functions even where to the abstract psychological thinker they have seemed nearly identical. There are no few elemental 'faculties' or forms of mental activity which work alike with any and every kind of content.

For instance, the correlation in adults between (1) memory for figures and (2) memory for unrelated words (memory being used to mean the power to keep a list in mind, after once hearing it, long enough to write it down) is not over .8; the correlation in pupils of the highest grammar grade between (1) quickness in thinking of the opposites of words and of the letters preceding given letters of the alphabet and (2) quickness in thinking of the sums of figures is not over .7. Yet the first pair of tests would commonly be used indiscriminately as tests of 'memory,' and the second pair as tests of 'association,' upon the supposition that the two members of each pair were practically identical traits. Even so apparently trivial a difference as that between drawing a line to equal a 100 mm. line and drawing a line to equal a 50 mm. line causes a reduction from perfect correlation. The resemblance is, for 37 young women students, only .77.

A table of the known degrees of relationship would abundantly confirm the statement that the mind must be regarded not as a functional unit, nor even as a collection of a few general faculties which work irrespective of particular material, but rather as a multitude of functions each of which involves content as well as form, and so is related closely to only a few of its fellows, to the others with greater and greater degrees of remoteness.

The mental sciences should at once rid themselves of the conception of the mind as a sort of machine, different parts of which sense, perceive, discriminate, imagine, remember, conceive, associate, reason about, desire, choose, form habits, attend to. Such a conception was adapted to the uses of writers of books on general method and arguments for formal discipline and barren descriptive psychologies, but such a mind nowhere exists. There is no one power of sense discrimination to be delicate or coarse, no capacity for uniform accuracy in judging the physical stimuli of the outside world. There are only the connections between sense-stimuli and our separate sensations and judgments thereof, some resulting in delicate judgments of difference, some resulting in coarse judgments. There is no one memory to hold in a uniformly tight or loose grip all the experiences of the past. There are only the particular connections between particular mental events and others, sometimes resulting in great surety of revival, sometimes in little. And so on through the list. Good reasoning power is but a general name for a host of particular capacities and incapacities, the general average of which seems to the namer to be above the general average in other individuals. Modern psychology has sloughed off the faculty psychology in its descriptions and analyses of mental life, but unfortunately reverts customarily to it when dealing with dynamic or functional relationships.

But it is just in the questions of mental dynamics and of the relationships of mental traits that we need to bear in mind the singularity and relative independence of every mental process, the thoroughgoing specialization of the mind. The mind is really but the sum total of an individual's feelings and acts, of the connections between outside events and his responses thereto, and of the possibilities of having such feelings, acts and connections. It is only for convenience that we call one man more learned than another instead of giving concrete lists of the information possessed by each and striking averages from all the particulars; that we call one man more rational than another instead of comparing two series of rational performances. In any one field the comparison may give a result widely different from the general average. So also with inhibition, concentration, or any other of the general names for forms of mental action.

It is easier to show that mental traits are not related in certain ways than it is to show directly just how they are related. There are three reasons for this. The first is that, until the work of Spearman in 1904, it was not known that any relation, calculated from other than complete measurements of the divergences to be related, would be reduced from its true amount toward 0. Wissler, for example, found ['01] the relation between accuracy in drawing a line to equal a 100 mm. line and accuracy in bisecting it, to be only .38, because he had as measures of the two traits to be related only the results with a single line. If each individual had been tested with hundreds of lines at different times so as to get his real total abilities the result would have been an *r* of .7 or .8. Since this fact of the 'attenuation' of the *r*'s obtained from imperfect original data was not realized at all until 1904, and is not realized by many students even now, the majority of the *r*'s that have been calculated are of doubtful meaning. For only very rarely have the measurements of individuals in any trait included many tests on many dates under many circumstances so as to give adequate measures. The true *r*'s can only be inferred from the obtained *r*'s by a guess at the degree to which the latter are attenuated. In the case just quoted the true *r* turns out twice the obtained; in other cases it turns out to be only a trifle greater, depending of course on the chance deviations of the original measures.

The second reason for our inability to give a clear, simple, positive account of the relations of mental divergences is that the divergences related have been measured in different groups and from all sorts of different central tendencies. The result is that many of the relations known are not directly comparable. For example *r* for general scholarship with quality of handwriting is, in children in the elementary school, being measured by their divergences from the central tendencies of their age and sex, about .3. *r* for general scholarship and desultory memory is, certain educated adults being measured by their divergences from the central tendencies of graduate students, also about .3. But the two relations are really not equal.

The third reason is that the fact itself would, in any case, be too complex for clear and simple positive description. The facts which we do know prove that if we knew all the facts the interrelations of mental traits would be irreducible to any easy representation. This is itself, however, an important fact about them.

Such a general positive description as can be given will be best given in graphic form. Let the series of horizontal lines in Fig. 89 be each a scale for some mental trait; let the right end always represent superiority in the trait; let the central tendency of adult human beings be in each case at the point on the scale where a vertical line down the center of the diagram would cut the scale. The C. T.'s for all traits are then in a

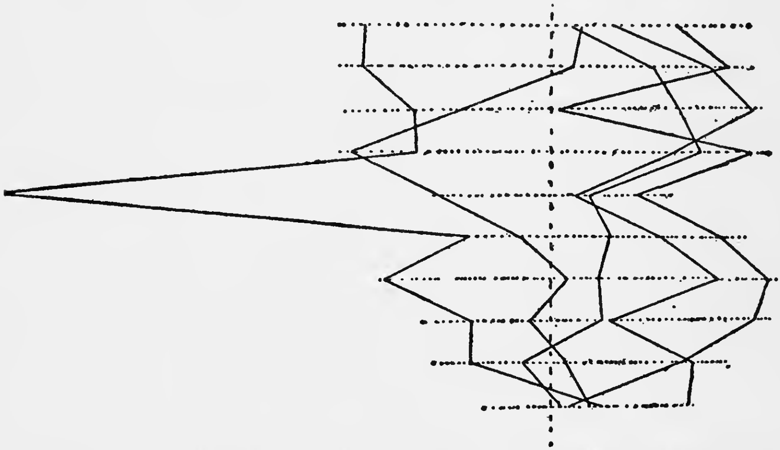


Fig. 89. Ten Traits Shown in Their Interrelations in the case of Each of Five Individuals. Three of the individuals are university teachers or graduate students. The other two are inferior day-laborers.

vertical line. Let a constant distance always equal the variability of the trait. Let any one individual be represented by a line joining all the points denoting each his ability in one trait. Five individuals are thus represented in Fig. 89, in ten traits. If now, instead of ten horizontal scales for the traits, there were the thousands necessary to inventory human faculty, and if, instead of five cross-lines for five individuals, we had thousands representing a fair sampling of all men, the picture would show a rough general parallelism of the cross-lines corresponding to the general tendency for efficiency in one trait to go with efficiency in others. There would be very few cross-lines from the extreme left of one scale to the extreme right of the other.

On the other hand there would be few horizontal lines cut by the same man's crosslines at the same distances from the C. T. There would be many cross-lines that departed from the general drift of the cross-lines a little; some that departed from it more; and a few that departed from it enormously.

The first impression would be of a general parallelism of the cross-lines, disturbed at haphazard, somewhat as geological strata are distorted from their parallelism in a mountain region. If, however, one examined the relations, he would find certain rough rules for explaining them and prophesying others. Other things being equal, two functions would be correlated in proportion as efficiency in them depends upon the status of the unknown "Common Factor," variously called General Intellect, Mental Energy, Ability to Learn, and the like. Perhaps, if we could define this factor, learn its symptoms, and test its strength in different individuals, we should find that it alone explained the main features of all the cross lines. Perhaps, on the other hand, the other relating forces would be even more important.

Amongst these other relating forces are the organization of the neurones into sensory and associative, their organization into more or less specific conduction-chains paralleling certain instincts, and the effects of certain common forms of training upon different functions to raise an individual's status in both or to raise his status in one at the expense of the other. Correlations seem to be closer within the analytical or abstracting functions than between these and others. So also within the purely mental associative functions like adding, completing words, giving opposites, and naming objects, than between one of them and one of the sensori-motor functions. The sensitivities seem to interrelate only loosely; and any one of them would relate very loosely to the associative or analytical functions, even when the latter was busied with data from that sense.

Certain functions quite diverse from the point of view of the classification into sensory, associative and analytic processes would be found to be related by reason of some instinctive tendency. Different degrees of the instinctive interest in persons might thus produce unexpectedly high correlations between love of ceremonies, ability in sociology and interest in literature. The variations in the love of mental activity are one root of many of the correlations between all sorts of efficiencies.

The circumstances of training would be seen to sometimes intensify and sometimes weaken original relations. Thus the variations in the length of school training, by connecting a long training in certain traits almost inevitably with a long training in certain other traits, make certain correlations more pronounced. The world over, there is a close relation between knowledge of Latin and knowledge of geometry, far beyond what natural interests would produce. A case of the weakening of an original relation by training might be furnished by ability in mathematics and ability in understanding human nature. The original relation is perhaps closer than the actual relation in adults, the common forms of training tending so often to develop one *or* the other. Roughly, the effect of the environment is to make closer the relations due to content. A mother comes to observe, attend to, and remember her children well; the plumber 'develops all his powers,' as the educational theorists were fond of saying, so far as plumbing goes, and neglects them beyond that point; the classical student takes a high position in every formal operation when Latin is the *datum* operated on.

Just what the original relations are will in the progress of research be discovered. It is unlikely that the relations of original capacities and instincts, including native interests, are so complicated as the relations amongst adult achievements and abilities. But present knowledge is insufficient to determine even the original relations.

CHAPTER XVI

The Nature and Amount of Individual Differences in Combinations of Traits: Types of Intellect and Character

Scientific studies of the natures of individual men in combinations of traits by the direct method have been very few and very inadequate in respect to both the number of individuals and the number of traits studied. The measurements of mental relations described in the last chapter are the main means of improving upon common opinion concerning the varieties of human nature in respect to total make-up, or to such large fractions of it as are roughly denoted by temperament, mode of thought, morals, mental health, manner of work, imagination, and the like.

A SAMPLE PROBLEM: INDIVIDUAL DIFFERENCES IN IMAGERY

As a sample of the problems and their treatment we may take the natures of individuals in respect to type of imagery, that is, in respect to the combination of:—vividness of visual images, fidelity of visual images, frequency of visual images, vividness of auditory images, fidelity of auditory images, and so on, through the list.

Early in the history of the scientific study of imagery it was noted that certain individuals were able to recall in memory presentations to one sense with a high degree of vividness and fidelity, but lacked this power in the case of presentations to some other sense. The existence of persons who, for instance, could get before the mind's eye vividly and with full detail a mental photograph, as it were, of a scene, but could not thus reproduce from within a melody, an itching nose, or a blow, naturally gave rise to the notion of the 'visualizing type.'

Such cases, of notable ability to get one sort of images and notable inability to get other sorts, were then carelessly assumed to be the rule. It was supposed that a high degree of vividness, fidelity and frequency in images from one sense tended to exclude an equally high degree in images from other senses. People were called visualizers, audiles, motiles, etc., with the meaning that the visualizers had more vivid, faithful and frequent visual images than other people and less vivid, faithful and frequent images from other senses, and similarly for the audiles, or motiles. In graphic form this view would give Fig. 90.

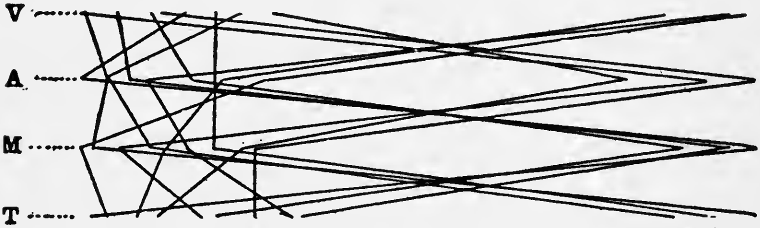


FIG. 90. The Interrelations of the Degree of Development of Visual, Auditory, Motor, and Touch Imagery according to the Theory of Pure Types. Imaginary horizontal lines at V, A, M and T are scales for the degree of vividness, fidelity and frequency of visual, auditory, motor and touch imagery respectively. The lowest degree is in each case at the left. 12 individuals are represented, each by a line crossing each of the scales at the point representing the individual's ability.

But the actual examination of individuals showed such exclusiveness or predominance of one sort of imagery to be the exception rather than the rule. To even superficial examination it was evident that human natures did not fit into the scheme of Fig. 90 at all well. Even those who believed unhesitatingly that human natures must be distributed around fairly distinct types in respect to imagery could not, try as they might, distribute individuals around these types. Meumann in fact admits that in all his studies of children he never found one such pure type. "How rare the pure types [of imagery] are amongst children is witnessed by the fact that in our extensive investigations of children at Zurich we have never found a perfectly pure type. Also I know of no case in the entire literature of the subject in which sure proof is given of the existence of a pure type in the case of children." ['07, I, p. 494] So new intermediate types, such as the auditory-motor, visual-motor, auditory-visual, or even visual-auditory-motor-intellectual (!) [Segal, '08], were introduced. There the matter remained until Betts ['09] actually measured a sufficient number of individuals in respect to the vividness and fidelity of non-verbal images from the different sense-fields, so that such cross-lines as those of Fig. 90 could be located by fact instead of by opinion.

The pillars of the doctrine were the separation of men into types according to the predominance of images from one sense, and the existence of inverse relations between the different sense-spheres in respect to the extent and perfection of imagery. Fact showed opinion to have been grossly in error as a result of its assumption that distinct types of some sort there must be. The contrary is true. Instead of distinct types, there is a continuous gradation. Instead of a few 'pure' types or many 'mixed' types, there is one type—mediocrity. Instead of antagonism between the development of imagery from one sense and that from other senses there is a close correlation. Fig. 91 is a fair sample of the facts found.

This case is instructive because the fate of many theories concerning distinct types of human nature in combinations of traits is likely to be the same as the fate of the doctrine of types of imagery according to the sense involved, with inverse relations between the development of imagery from one sense-field and that from other fields. In the case of temperament, for example, we have the same history. Extreme cases are given names and made into types. Verbal contrasts are supposed to have real existence. Supplementary types are invented to help out the discrepancies between the imagined types and the real distribution of individuals. And it is highly probable that, when actual measurements are made, mediocrity—a temperament moderately sanguine, choleric, phlegmatic, and melancholy; moderately slow, quick, shallow, intense, narrow and broad; moderately slow-shallow, slow-intense-narrow; moderately *everything*,—will be found to be the one real type.

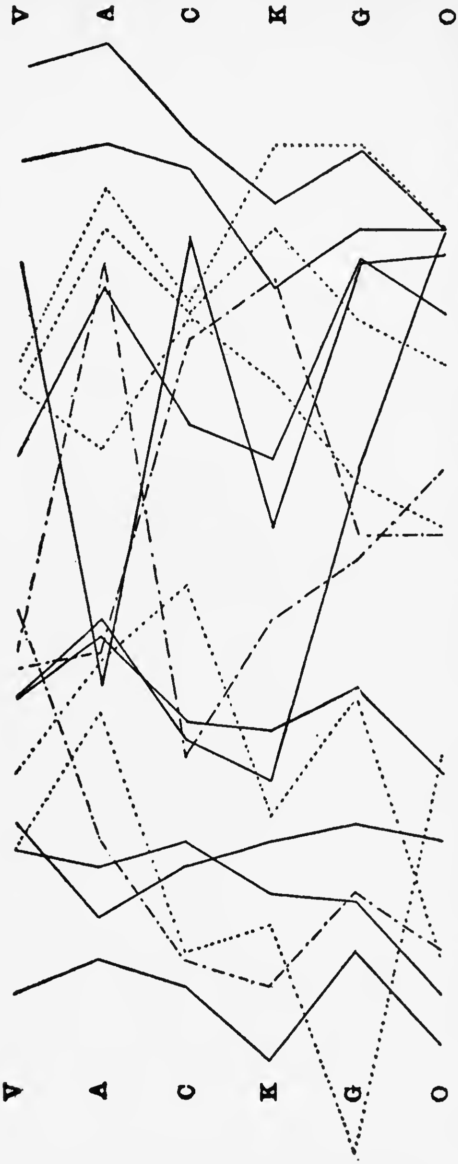


Fig. 91. The Interrelations of Visual, Auditory, Cutaneous and Other Forms of Imagery as Found by Betts in the case of 17 Graduate Students or Teachers of Psychology. The horizontal lines VV, AA, CC, KK, GG and OO are scales for the vividness, fidelity and frequency of visual, auditory, cutaneous, kinaesthetic, gustatory and olfactory images voluntarily summoned. The lowest degree is in each case at the left. Each individual is represented by a cross-line as in Figs. 89 and 90. The differences in the drawing of the cross-lines (line, dot, dot and dash) have no significance, being intended simply to aid the eye in realizing the status of any individual.

THE THEORY OF MULTIPLE TYPES AND THE SINGLE-TYPE THEORY

The sample problem shows well two extreme views which may be taken of the varieties of human natures, of the same sex, race and degree of maturity, in respect to any combination of traits. On the one hand is the theory of multiple types, a theory which separates men more or less sharply into classes, and describes a man by naming the class to which he belongs. On the other hand is the theory of a single human type, a theory which joins all men one to another in a continuity of variation and describes a man by stating the nature and amount of his divergences from the single type.

By the theory of a single type, one make-up can be conceived such that from it all individuals would differ less than they would from any other one make-up, and such that, the greater the divergences, the rarer they would be. By the theory of multiple types, no such single true central tendency would exist. By the theory of multiple types, if a number, K, of 'typical' natures or make-ups are most favorably taken and if divergences of all individuals are measured from, in each case, that nature which the individual most resembles, the total sum of divergences is enormously reduced below what it would be if they had been measured all from some one nature. By the theory of a single type, this reduction in the sum of divergences due to measuring each individual's divergence from any one of K natures, is much less.

These two doctrines can be made clear by graphic illustrations. Let the amount of each trait in the combination be scaled, as by our custom, horizontally, the center always representing the mode. Let the nature or make-up of each individual be represented by the points where a cross-line denoting him cuts the scale lines. The theory of multiple types then gives something like Fig. 92, and that of one type something like Fig. 93. All the cross-lines of Fig. 92 can be represented as minor divergences from five typical cross-lines, far better than can all the cross-lines of Fig. 93. Those of Fig. 93 can be far better represented by one typical cross-line than can those of Fig. 92.

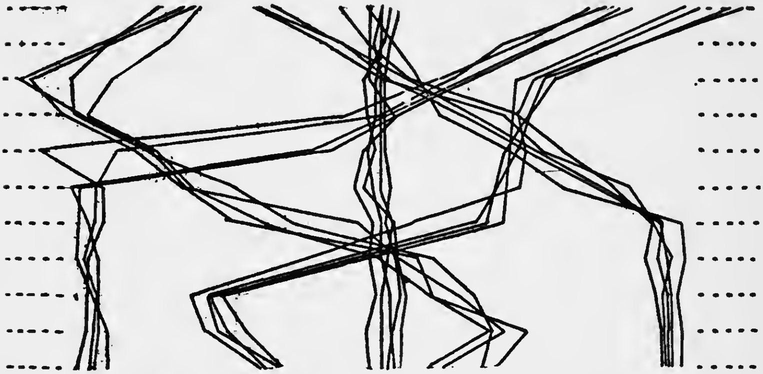


FIG. 92. A Graphic Representation of the Multiple-type Theory in the case of Combinations of Traits. The 11 horizontal dotted lines (drawn only at the extremes) represent scales for 11 traits. Each cross-line represents, by its location, the amounts of the 11 traits in one individual.

A radically different scheme for graphic representation will enable the reader to think of the meaning of the two theories for a very large combination of traits, say, total mental nature.

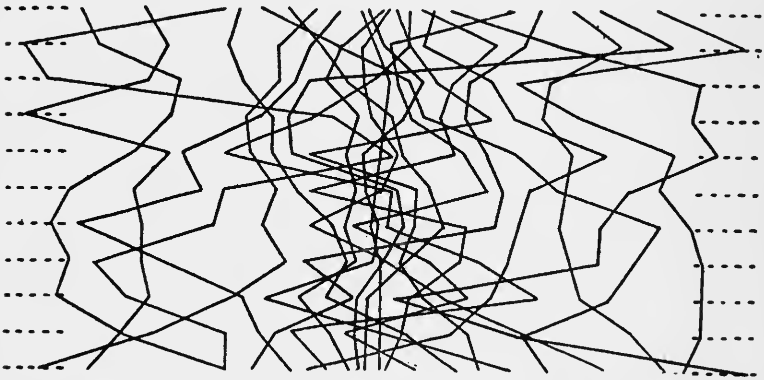


FIG. 93. A Graphic Representation of the Single-type Theory in the case of Combinations of Traits. The 11 horizontal lines represent scales for 11 traits. Each cross-line represents the amounts of the 11 traits in one individual.

Suppose that each trait is scaled into 5 amounts, 1, 2, 3, 4 and 5. Suppose that there are 5 traits, a, b, c, d and e, concerned in the total complex. Suppose 3 to be the mode in each trait.

Consider every possible cross-line, every possible series of amounts in the 5 traits that can be possessed by an individual (but without fractionizing any scale farther than into fifths). There will be 3125 such series possible (5 to the 5th power). Now suppose each of these 3125 series of amounts,—each of these 'natures,'—to be represented by a spot of, say, .01 sq. in. on a surface. Suppose further that the center spot of the map represents the series 33333 (the traits being now and later always taken in order a, b, c, d and e). Suppose that around this center spot and adjoining it are placed all the spots or areas representing those series diverging from the 33333 series by only one unit, *e. g.* 23333, 43333, 32333, etc. Suppose that the next ring is made up of the spots or areas representing those series diverging from the modal series by two units, *e. g.* 22333, 43433, 53333, etc. Continue till, in the outermost ring, we have areas representing the series 55555, 11111, 11555, 51115, etc., which have the greatest possible divergence from the modal series.

Now suppose that the relative frequencies with which these series appear in men are represented each by the height of a column erected with the appropriate area as base. Consider the appearance of the resulting relief-map according to the two theories. By the single type theory, it would be highest over some one spot—in our illustration, over the center or 33333 spot.* By the multiple type theory, it would have several maxima over several spots. By the single type theory, the average height would be less and less over each successively more distant concentric ring. The whole relief map would look like a mountain, with, possibly, many radiating and cross ridges, casual valleys and eminences, but with, always, a decreasing elevation, the larger the radius of a circle drawn about the summit. By the multiple type theory, the average elevation need bear no relation to the distance from the center spot; the

*This would be the case in the special form that the theory would almost certainly take in the case of combinations of 5 traits, each unimodal, with the mode at mediocrity.

depressions would be between the several peaks each representing a type. In proportion as the types were held to be very sharply distinct and widely separated, these intervalleys would be deep, even to the original surface, and wide.

Extend now the process of construction of these models to thousands of traits each scaled in fifty or a hundred amounts, and one has the contrast of the multiple and single type theories applied to men's total natures.

It is not necessary to try to decide between these two theories, or to determine just what compromise is the true one. It is better to accept frankly our ignorance of just how individuals do differ in combinations of traits until they have been measured in respect to all the traits involved.

Since, however, many writers about human nature openly or tacitly assume the truth of the multiple type theory in a pronounced form, and are governed by it in their methods of research, of interpretation and of practical control, it will be useful to consider briefly some of the arguments in favor of the single type theory.

The first is the fact that, in proportion as exact measurements have been applied, evidence expected to favor the multiple type theory has turned out in favor of the single type theory. It is true that such cases are very rare, and that, until they are much increased in number, little should be inferred from them. But the fact remains that the single type theory arose from exact measurements, while its opposite came from speculative prepossessions.

The second is the rarity of the inverse correlations between desirable traits upon which so many of the supposed multiple types are based. We know that eye-minded and ear-minded, quick and careful, broad and deep, sensorial and intellectual, men of thought and men of action, and the like do not really represent human nature's varieties in the combinations referred to. If two horizontal scales are drawn for 'ability to learn through the eye' and 'ability to learn through the ear,' and the crosslines are drawn for a thousand individuals, they will not go as in Fig. 94 but as in Fig. 95. So also for scales for quantity of work and quality of work, and so on through the list.

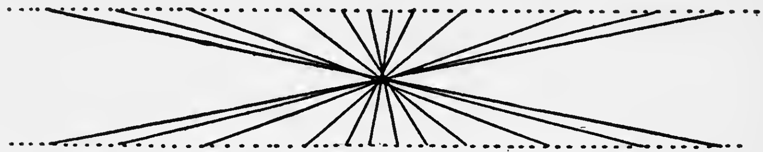


FIG. 94. A graphic representation of the condition of individuals in a combination of two traits, if these are very antagonistic or inversely correlated. The general scheme of the diagram is that used in FIGS. 89 to 93.

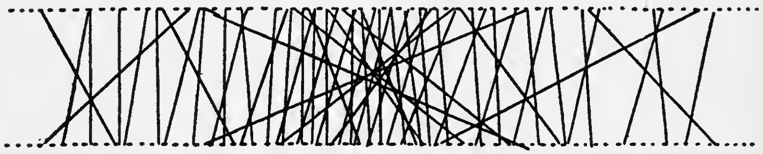


FIG. 95. A graphic representation of the condition of individuals in a combination of two traits, if these are closely and positively related.

The third is the fact that the single traits involved are so often distributed each approximately symmetrically around one mode and that their intercorrelations are so often approximately rectilinear. In so far as such is the case, no matter what the closeness of the correlations may be, the distribution of the individuals in respect to the combination of traits will be around one mode or peak with less and less frequency the more they diverge from that mode. The mode for the combination will be a nature which is at the mode in all the single traits.

The fourth is the fact that investigators who are strongly in favor of the multiple type theory and accustomed to interpret facts in harmony with it, yet find so few actual cases of it. Meumann, for instance, ['07, vol. I] clearly accepts the theory *in general* and demands that educational practice should give much attention to the classification of pupils under distinct types. But in concrete particulars he rarely illustrates it.

He says ['07, vol. I, pp. 331-332]: "By establishing types we orient ourselves in the endless possibilities of individual differences, and if we can place an individual under a type

. . . . in any respect we thereby have pointed out a group of universal characters in his mental life, which he in general shares with some individuals and by which he is in general distinguished from others." But he does not establish such types. The majority of the differences which he does report as 'typical' are differences between two extremes of the same trait. Intermediate conditions are in some of these cases demonstrably, and in all cases probably, more typical than the supposed types. And this, indeed, Meumann, in some cases, admits.

Lastly, I may mention the fact that satisfactory proof of the existence of a distribution of human individuals after the fashion demanded by the multiple type theory has never been given in a single case, and that the evidence offered by even the most scientific of the theory's adherents is such as they would certainly themselves consider very weak if they were not already certain that types of some sort there must be. Thus a fair-minded perusal of Stern's *Psychologie der Individuellen Differenzen,* designed to be a description of the types into which human nature falls, is an almost sure means of stimulating a shrewd student to the suspicion that intermediate conditions are more frequent than the supposed types, and that there are far more simply ordinary people than there are of all the 'types' put together. I report, to illustrate this point, the substance of all that Stern says in favor of the multiple type theory in his five chapters on Sensitivity, Perception, Memory, Association and Apprehension ['00, pp. 40-77].

On p. 43 f. he says, "We know the enormous gap which exists between the unmusical and the musician in the discrimination of pitch, between the perfumer and the ordinary person in the recognition of odors, between the painter and the bookworm in the delicacy of color perception." On the contrary, between the keenest of the non-musical and the dullest of musicians in the discrimination of pitch there is no enormous gap, but an enormous overlapping (see Spearman, '04 b, pp. 90 and 92). Stern himself later points out that a little special practice bridges the 'enormous gap.'

Stern mentions (p. 46) "the types of the external observer (the experimental scientist, possibly) and of the introspective thinker (the mathematician or metaphysician, possibly)." But these are not distinct, contrasting types. The experimental scientist is far more likely to be a good mathematician than is the ordinary man. Mathematical ability and interest are in no sense confined to the metaphysicians. The good external observer may be excellent at introspection, and the man with a strong interest in his inner life of thought is much more likely than the average man to have a strong interest in external facts.

Stern's next group of types are the 'Anschauungstypen.' Individuals of the visual type, for instance, "imagine and dream in the most vivid visual images; they notice and retain colors, forms, faces with especial ease; they reproduce what is spoken predominantly with the aid of images of the printed words; in fact they in general construct their world of ideas in great measure out of visual elements." The more such extreme development of any one sense-sphere exists in a man, the more "the others are restricted to their necessary, indispensable and irreplacable functions." (p. 48) But here again the facts are in opposition. Such cases as Stern describes are extreme and very rare divergences from common mediocrity, not centers near which most men are located.

Stern further divides men into a 'formal' and a 'material' or 'content' type in their 'Anschauung' or mental dealings with concrete objects. The formal type gets the rhythm of a piece of music easily, but the melody only with difficulty. The 'content' type gets the melody well, but the rhythm badly. "The former type notices and retains above all the temporal grouping of the sounds, whose qualitative nature and relations stay far in the background. With the latter type, on the contrary, it is just the material of the sensations, the succession of different pitches, which is made the instrument of musical perception." (p. 55) But no evidence is offered of this supposed inverse correlation, and all that is known of the relation leads to the expectation that it is positive, the individual best at getting rhythm being better than the average at getting melody also.

Stern supposes these contrasting types to exist for spatial facts also. Ability with position and outline denotes one type; ability with color, light and shade, another type (p. 57). But here again the supposed inverse relation is really direct. The supposed types are two extreme and rare divergences from the real single type, an individual who is mediocrely gifted with both the 'form' and the 'material,' with both line and color.

In the case of memory, Stern reports, not multiple types, but gradations from single types, with the single exception that he quotes with modified approval, "Who learns easily, forgets easily also; who learns with difficulty, retains better." This is, again, contrary to the facts.

Under types of apprehension or apperception, is given the classification of individuals as describers, observers, emotional, and scholarly, made by Binet ['97] on the basis of their written descriptions of a single picture, and their classification as describers, observers, imaginative and poetic, and scholarly, made (also by Binet) on the basis of written descriptions of a cigarette. Stern points out that much more elaborate examinations of the children are necessary. Leclère ['98] got, in a single experiment, seven types, which, also, are not at all distinct. Of course, many samples of each child's 'apprehensions' or 'apperceptions' of objects must be collected and classified independently by many competent psychologists, and the exact rating of each person on scales for scholarliness, imaginativeness, and so on must be given, before one can tell whether the individuals are grouped around one type or around several. The question is not whether Professor Binet can pick out certain papers that are distinctly different, but whether the common judgment of experts will rank the individuals in classes that are distinctly different.

The above is all that is given as evidence for the multiple type theory in the first thirty-five pages of Stern's inventory of types. The balance of his discussion, devoted to Types of

Attentiveness, of Activity in Combining Facts, of Judgment, of Reaction, of Feeling, of the Tempo of Mental Life and of Mental Work, would give, if anything, less satisfactory evidence. The same lack would be found in the writings of Kraepelin, Binet, Meumann and their students.

In Stern's recent general statement of the multiple-type theory ['11, p. 168 ff.], it is supported by no better defenses. The theory itself is here stated more cautiously, no great distinctness of separation of one type from another being asserted. "Between one type and its neighboring types the boundaries are always fluctuating; the transitional forms, in whose case it is questionable whether they belong to one or the other type or perhaps to an intermediate form, are not occasional abnormalities but necessary features of the organization of human nature into types." ['11, p. 173]

In general the case is stronger for the single type theory in combinations of original traits than in traits produced by training. One form of training—one environment—may possibly produce in every individual who is subjected to it certain large increases in certain mental traits and certain large decreases in others. All its subjects then tend thereby to differ much from all other men, and to be much alike among themselves, that is, to cluster around a type of their own. Such is not without reason supposed to be the case with residence in a given nation, frontier life, city *versus* country life, household industry *versus* factory industry, slavery *versus* freedom, and the like. Just as, in any one trait, the action of disproportionately large factors may cause multimodality, so, in any combination of traits, the action of disproportionately large factors may cause multitypality. And doubtless there would appear many minor eminences upon our relief-map of human nature due to the influence of 'nation lived in,' 'language spoken,' 'occupation followed,' and the like. The form of distribution of individuals in respect to combinations of traits, as in respect to single traits, is a secondary result of the nature and interrelations of the factors that produce the amounts of the traits. There is no mysterious force of mental life herding all individuals of the same sex, age and race around mediocrity. Nor is there such a force separating them out into clusters around distinct types. The former condition certainly occurs in the case of many combinations of traits. The latter may very well occur in the case of others. Knowledge of just what does occur demands objective measurements of each of many individuals in each of the traits in each of the combinations in question.

INDIVIDUAL DIFFERENCES IN THE AVERAGE AMOUNT OF A COMBINATION OF TRAITS

There are many combinations of traits which can be reduced to single traits by abstraction from some of their particulars. Suppose, for example, that A and B are measured in respect to efficiency in marking A's, in marking words containing each the two letters a and t, in marking hexagons on a sheet of various simple geometrical forms, in marking grays of a certain intensity on a sheet with 200 squares of grays of five intensities, and in marking misspelled words on a sheet containing a passage with 100 out of 500 words misspelled. Suppose the results to be, in terms of the variability for each trait:—

	A	B
Marking A's	—1.1	+1.0
" a — t words	—1.4	+0.7
" hexagons	—.6	+1.2
" grays	.0	—.2
" misspelled words	—1.7	+1.4

B — A equals 2.1, 2.1, 1.8, —.2 and 3.1 respectively.

If we abstract from the particular differences and ask only concerning the condition of A and B, and their difference, in *average efficiency in marking these five sorts of visual objects,* the result is that A= —.96 (—4.8 divided by 5), B=+.82 (+4.1 divided by 5), and B — A = 1.78.

Such abstraction from certain particulars of each of a combination of traits can be, and is, in both ordinary and scientific thinking, carried so far as to unite in a single trait very diverse features of intellect and character. From the combination of all the accuracies of discrimination with this and that length, color, weight and the like, may be got the one trait, *accuracy in sensory discrimination.* From the quickness of formation of each of a thousand habits, is derived the single trait, *rate of learning.* Accuracy, quickness, efficiency, permanence, amount of improvement, rate of improvement, and acceleration or retardation in the rate of improvement, are important cases of the measurement on one scale of some feature of an individual's condition in a group of traits. Originality, courage, timidity, suggestibility, scholarship, judgment, interest and curiosity are samples from a long list that could be made of terms, each used with comparatives to denote, though very crudely, a man's position on a single scale. This position or amount would, however, be the resultant of many manifestations of what would have to be scored as a combination of many traits, if represented in full, concrete detail.

For all such one-scale representations of combinations of traits, the entire theory of single traits given in Chapter XIV holds good. In particular, the single type theory holds of them with fewer exceptions. For some one large cause will much less often act upon a man with the same effect in all the traits of a combination than in some one of them. So, whereas, in discrimination of the tastes of wines or teas or the like, men may be divided into an ordinary and an expert class, they will not be, in respect to accuracy of sensory discrimination in general. Similarly, though, in knowledge of the Latin language, men may fall into two groups,—an ignorant group and a group varying around some knowledge,—in knowledge of languages in general, they do not.

From the discussion of the previous paragraph it is evident that, unless peculiar causative or selective agencies are at work, there will be a few individuals who will possess so little of any given capacity or quality, or combination of capacities or qualities, as to be obviously 'defectives' in it, as well as a few who will possess so much as to be obviously 'prodigies.' There will be a larger number who will possess so little as to merit the popular term 'weak' in color vision, memory, self-control, moral sense, general intelligence or whatever the trait may be. These, again, will be balanced by an approximately equal number of 'remarkable' or 'exceptionally gifted.'

If the mental trait in question is the compound of many traits which we call intelligence, we shall find at the lower end of the distribution curve children whom medical diagnosis would name idiots, and next them a number who would be termed imbeciles, and nearer still to the average the group to whom the name weak- or feeble-minded would be applied. If the mental trait is the compound called 'morality,' the individuals at the low extreme will perhaps be diagnosed as cases of 'moral insanity' or as 'moral degenerates.' If the trait be more specific, for instance if it be ability to learn to spell, ability to learn to read, cruelty, musical ability, memory, visualizing power or what not, we shall find few, if any, special names for different degrees of its possession, though there will as truly be defectives in respect to any such specific mental trait as in respect to general intelligence.

The means which educational endeavor will use and the results which may be expected therefrom will, in the case of any individual, depend upon his amount of the trait in question. No one, unless he were himself an 'idiot' in the trait of common sense, would train a genius and an idiot alike or expect them to develop alike. At present there is a widespread practice of providing separate treatment at home or in institutions for idiots and imbeciles, though some are to be found in the common schools. And there is a growing demand for institutions and separate classes for the feeble-minded. Notable moral defectives are being cared for in separate classes in some cities. They also, when the parents are wealthy, find refuge in private schools of a certain type and in the somewhat mercenary ministrations of private tutors. The children who are extremely divergent in their great superiority to the average are not systematically given any special attention except here and there by systems of rapid promotion.

For the proper treatment of children who diverge extremely from the type we need knowledge of the exact distribution of all the mental traits which we desire to develop or abolish, of the causes which determine an individual's station in each, of the symptoms by which we may conveniently find out any one's station in each, and of the agencies, educational, hygienic and medicinal, which alleviate or intensify the different conditions. The last involves the study of the differential action of stimuli upon individuals of different stations.

It is not within the province of this volume to describe the concrete details of the almost infinite variety of extreme human natures, or to criticize means of educational diagnosis and treatment. The general principle that each child is understood only when his exact station in every trait of consequence is known, and is properly educated only when means are chosen in view of his individual make-up, is an obvious consequence of the last three chapters. The student who proceeds to the study of extreme degrees of intellect, morality, or skill will have had an adequate general introduction in the principles of human variability in general of which these extreme degrees are one consequence.

Bibliography of References Made in the Text

Abelson, A. R........'08 *Fatigue and Its Measurement by the Aesthesiometer.* International Magazine of School Hygiene, vol. 5, pp. 347-487.

Abelson, A. R........'11 *The Measurement of Mental Ability of Backward Children.* British Journal of Psychology, vol. 4, pp. 268-314.

Amberg, E...........'95 *Ueber den Einfluss von Arbeitspausen auf die geistige Leistungsfähigkeit.* Psychologische Arbeiten, vol. 1, pp. 300-377.

Arai, T..............'12 *Mental Fatigue.* Teachers College, Columbia University Contributions to Education, No. 54.

Aschaffenburg, G.'96, '99 *Experimentelle Studien über Associationen.* Psychologische Arbeiten, vol. 1, pp. 209-299, and vol. 2, pp. 1-83.

Baur, A..............'02 *Die Ermüdung der Schüler in neuem Lichte.*

Bergström, J. A.....'94a *An Experimental Study of Some of the Conditions of Mental Activity.* A. J. P., vol. 6, pp. 247-274.

Bettman, S...........'95 *Ueber die Beeinflussung einfacher psychische Vorgänge durch körperliche und geistige Arbeit.* Psychologische Arbeiten, vol. 1, pp. 152-208.

Betts, G. H..........'09 *The Distribution and Functions of Mental Imagery.* Teachers College, Columbia University Contributions to Education, No. 26.

Binet, A..............'97 *Psychologie individuelle: La description d'un objet,* L'année psychologique, vol. 3, pp. 296-332; and *La description d'une cigarette,* Revue de Psychiatrie, 1897, pp. 235-243.

Binet, A. and Henri, V.............'98 *La fatigue intellectuelle.*

Binet, A..............'05 *Recherches sur la fatigue intellectuelle scolaire et la mesure qui peut en être faite au moyen de l'esthésiomètre.* L'année psychologique, vol. 11, pp. 1-37.

Blazek, B ....'99 *Ermüdungsmessungen mit dem Federaesthesiometer an Schülern des Franz-Joseph-Gymnasiums in Lemberg.* Zeitschrift für Pädagogische Psychologie, vol. 1, pp. 311-325.

Bolton, T. L...........'92 *The Growth of Memory in School Children.* A. J. P., vol. 4, pp. 362-380.

Bolton, T. L...........'02 *Ueber die Beziehungen zwischen Ermüdung, Raumsinn der Haut und Muskelleistung.* Psychologische Arbeiten, vol. 4, pp. 175-234.

Bonser, F. G.........'10 *The Reasoning Ability of Children of the Fourth, Fifth, and Sixth School Grades.* Teachers College, Columbia University Contributions to Education, No. 37.

Book, W. F...........'08 *The Psychology of Skill: with Special Reference to Its Acquisition in Typewriting.* University of Montana Publications in Psychology: Bulletin No. 53, Psychological Series No. 1.

Bourdon, B...........'95 *Observations comparatives sur la reconnaissance, la discrimination et l'association.* Revue philosophique, vol. 40, pp. 153-185.

Brown, W.............'09 *Some Experimental Results in Correlation.* Comptes Rendus du VIme Congrès International de Psychologie, p. 571 ff.

Brown, W.............'10 *Some Experimental Results in the Correlation of Mental Abilities.* British Journal of Psychology, vol. 3, pp. 296-322.

Bruner, F. G.........'08 *The Hearing of Primitive Peoples.* Archives of Psychology, No. 11.

Burgerstein, L........'91 *Die Arbeitscurve einer Schulstunde.* Zeitschrift für Schulgesundheitspflege, vol. 4, pp. 543-562 and 607-627.

Burris, W. P........'03 *The Correlations of the Abilities Involved in Secondary School Work,* Columbia Contributions to Phil., Psy. and Ed., vol. 11, No. 2, pp. 16-28.

Burt, C...............'09 *Experimental Tests of General Intelligence,* British Journal of Psychology, vol. 3, pp. 94-177.

Cattell, J. McK. and
Dolley, C. S. ....'95 *On Reaction Times and the Velocity of the Nervous Impulse.* Memoirs of the Nat. Acad. of Sci., vol. 7, pp. 391-415.

Cattell, J. McK. ....'03 *A Statistical Study of Eminent Men,* Popular Science Monthly, vol. 62, pp. 359-377.

Cattell, J. McK. ....'06 *A Statistical Study of American Men of Science, III. The Distribution of American Men of Science.* Science, New Series, vol. 24, pp. 732-742.

Cornman, O. P. ....'02 *Spelling in the Elementary School; An Experimental and Statistical Investigation.*

Courtis, S. A. ....'09 *Measurement of Growth and Efficiency in Arithmetic,* Elementary School Teacher, vol. 10, pp. 58-74, 177-199.

Courtis, S. A. .....'11, '12 *Report on the Courtis Tests in Arithmetic.* In the Interim Report of the Committee on School Inquiry of the Board of Estimate and Apportionment of the City of New York.

Davenport, C. B. ......'10 *Eugenics.*

De Candolle, A. ....'73 (Second Edition, '85), *Histoire des sciences et des savants depuis deux siècles, etc.*

Dodge, R. ....'13 *Mental Work: A Study in Psycho Dynamics.* Psy. Rev., vol. 20, pp. 1-42.

Dolley, C. S., and
Cattell, J. McK. ....'95 *On Reaction Times and the Velocity of the Nervous Impulse.* Memoirs of the Nat. Acad. of Sci., vol. 7, pp. 391-415.

Donovan, M. E., and
Thorndike, E. L. ......'13 *Improvement in a Practice Experiment under School Conditions.* A. J. P., vol. 24, pp. 426-428.

Earle, E. L. ....'03 *The Inheritance of the Ability to Learn to Spell.* Columbia Contributions to Phil., Psy. and Ed., vol. 11, No. 2, pp. 41-44.

Ebbinghaus, H. ....'97 *Ueber eine neue Methode zur Prüfung geistiger Fähigkeiten und ihre Anwendung bei Schulkindern.* Zeitschrift für Psychologie, vol. 13, pp. 401-459.

Ellis, A. C., and
Shipe, M. M. ....'03 *A Study of the Accuracy of the Present Methods of Testing Fatigue.* A. J. P., vol. 14, pp. 232-245.

Ellis, H. ....'94, '04 (Fourth edition, '04) *Man and Woman.*

Ellis, H. ....'04 *Study of British Genius.*

Fay, E. A. ....'98 *Marriages of the Deaf in America.*

Friedrich, J. ....'97 *Untersuchungen über die Einflüsse der Arbeitsdauer und der Arbeitspausen auf die geistige Leistungsfähigkeit der Schulkinder.* Zeitschrift für Psychologie, vol. 13, pp. 1-53.

Galton, F. ....'69, '92 *Hereditary Genius: An Inquiry into its Laws and Consequences* (first edition '69, second edition '92). References are to second edition.

Galton, F. ....'74 *English Men of Science: Then Nature and Nurture.*

Galton, F. ....'83 *Inquiries into Human Faculty and its Development.*

Galton, F. ....'89 *Natural Inheritance.*

Germann, G. B. ....'99 *On the Validity of the Aesthesiometric Method as a Measure of Mental Fatigue.* Psy. Rev., vol. 6, pp. 599-605.

Gilbert, J. A. ....'94 *Researches on the Mental and Physical Development of School Children.* Studies from the Yale Psychological Laboratory, vol. 2, pp. 40-100.

Goddard, H. H. ....'12 *The Kallikak Family.*

Goldmark, J. C. ....'12 *Fatigue and Efficiency.*

Griesbach, H. ....'95 *Energetik und Hygiene des Nervensystems.* (Reprinted from vol. 24 of the Archiv für Hygiene.)

Griesbach, H. ....'05 *Weitere Untersuchungen über Beziehungen zwischen geistiger Ermüdung und Haut-Sensibilität.* International Magazine of School Hygiene, vol. 1, pp. 317-417.

Hahn, H. H., and
Thorndike, E. L. ......'14 *Some Results of Practice in Addition Under School Conditions.* Journal of Educational Psychology, vol. 5, No. 2.

Hart, B., and Spearman, C. ....'12 *General Ability, Its Existence and Nature.* British Journal of Psychology, vol. 5, pp. 51-84.

Heck, W. H. ....'13 *A Study of Mental Fatigue.*

Heck, W. H. ....'13 *A Second Study of Mental Fatigue in Relation to the Daily School Program.* Psychological Clinic, vol. 7, pp. 29-34.

Henri, V., and Binet, A. ....'98 *La fatigue intellectuelle.*

Hëuman, G. ....'04 *Ueber die Beziehungen zwischen Arbeitsdauer und Pausenwirkung.* Psychologische Arbeiten, vol. 4, pp. 538-602.

Heymans, G., and Wiersma, E. ......'06 (a) *Beiträge zur speziellen Psychologie auf Grund einer Massenuntersuchung.* Zeitschrift für Psychologie, vol. 42, pp. 81-127, 258-301.

Heymans, G., and Wiersma, E. ......'06 (b) *Beiträge zur speziellen Psychologie auf Grund einer Massenuntersuchung.* Zeitschrift für Psychologie, vol. 43, pp. 321-373.

Heymans, G., and Wiersma, E. ....'07 *Beiträge zur speziellen Psychologie auf Grund einer Massenuntersuchung.* Zeitschrift für Psychologie, vol. 45, pp. 1-42.

Heymans, G., and Wiersma, E. ....'08 *Beiträge zur speziellen Psychologie auf Grund einer Massenuntersuchung.* Zeitschrift für Psychologie, vol. 46, pp. 321-333.

Hoch, A., and Kraepelin, E. ....'95 *Ueber die Wirkung der Theebestandtheile auf körperliche und geistige Arbeit.* Psychologische Arbeiten, vol. 1, pp. 378-488.

Holmes, M. E. ....'95 *The Fatigue of a School Hour.* Ped. Sem., vol. 3, pp. 213-234.

Höpfner, L. ....'94 *Ueber die geistige Ermüdung von Schulkindern.* Zeitschrift für Psychologie, vol. 6, pp. 191-229.

Hylan, J. P., and Kraepelin, E. ....'02 *Ueber die Wirkung kurzer Arbeitszeiten.* Psychologische Arbeiten, vol. 4, pp. 454-494.

Jacoby, P............'81 (Second edition, '04) *Etudes sur la sélection dans ses rapports avec l'hérédité chez l'homme.*

Kafemann, R.........'02 *Ueber die Beeinflussung geistiger Leistungen durch Behinderung der Nasenathmung.* Psychologische Arbeiten, vol. 4, pp. 435-454.

Kemsies, F...........'98 *Arbeitshygiene der Schule auf Grund von Ermüdungsmessungen.*

King, A. C..........'07 *The Daily Program in Elementary Schools.* (An unpublished study, on file in the library of Teachers College, Columbia University.)

Kirby, T. J...........'13 *Practice in the Case of School Children.* Teachers College, Columbia University Contributions to Education, No. 58.

Kraepelin E.,, and Hoch, A.............'95 *Ueber die Wirkung der Theebestandthiele auf körperliche und geistige Arbeit.* Psychologische Arbeiten, vol. 1, pp. 378-488.

Kraepelin, E., and Hylan, J. P..........'02 *Ueber die Wirkung kurzer Arbeitszeiten.* Psychologische Arbeiten, vol. 4, pp. 454-494.

Kraepelin, E., and Rivers, W. H. R......'96 *Ueber Ermüdung und Erholung.* Psychologische Arbeiten, vol. 1, pp. 378-488.

Kraepelin, E.........'02 *Die Arbeitscurve.* Philosophische Studien, vol. 19, pp. 459-507.

Kraepelin, E..........'03 *Ueber Ermüdungsmessungen.* Archiv für die gesamte Psychologie, vol. 1, pp. 9-30.

Krueger, F., and Spearman, C..........'06 *Die Korrelation zwischen verschiedenen geistigen Leistungsfähigkeiten.* Zeitschrift für Psychologie, vol. 44, pp. 50-114.

Larguier des Bancels, J.'99 *Essai de comparison sur les différentes methodes proposées pour la mesure de la fatigue intellectuelle.* Année psychologique, vol. 5, pp. 190-201.

Laser, H.............'94 *Ueber geistige Ermüdung beim Schulunterricht.* Zeitschrift für Schulgesundheitspflege, vol. 7, pp. 2-22.

Learoyd, M. W.......'96 *The Continued Story.* A. J. P., vol. 7, pp. 86-90.

Leclère, A...........'98 *Description d'un objet.* L'année psychologique, vol. 4, pp. 379-389.

Leuba, J. H..........'99 *On the Validity of the Griesbach Method of Determining Fatigue.* Psy. Rev., vol. 6, pp. 573-598.

Lindley, E. H........'00 *Ueber Arbeit und Ruhe.* Psychologische Arbeiten, vol. 3, pp. 482-534.

Lobsien, M...........'99 *Unterricht und Ermüdung: Experimentelle Untersuchungen.* Pädagogisches Magazin, Heft. 127.

Loewenfeld, R........'03 *Ueber die geniale Geistesthätigkeit.*

Marsh, H. D..........'06 *The Diurnal Course of Efficiency.* Archives of Phil., Psy. and Sci. Methods, No. 7.

Mayo, M. J..........'13 *The Mental Capacity of the American Negro.* Archives of Psychology, No. 28.

Meumann, E.........'07 *Vorlesungen zur Einführung in die experimentelle Pädagogik und ihre psychologischen Grundlagen.*

Miesemer, K..........'02 *Ueber psychische Wirkungen körperlichen und geistiger Arbeit.* Psychologische Arbeiten, vol. 4, pp. 375-434.

Mosso, A............... *Fatigue.*

Noikow, P. M.......'08 *Aesthesiometrische Ermüdungsmessungen.* International Magazine of School Hygiene, vol. 4, pp. 437-481.

Odin, A..............'95 *Genèse des grands hommes, gens de lettres français modernes.*

Oehrn, A............'95 *Experimentelle Studien zur Individual-psychologie.* Psychologische Arbeiten, vol. 1, pp. 92-151.

Orr, F. I., and
Rosanoff, A. J.....'11 *A Study of Heredity in Insanity in the Light of the Mendelian Theory.* American Journal of Insanity, vol. 68, pp. 221-261.

Pearson, K...........'04 *On the Laws of Inheritance in Man, II. On the Inheritance of the Mental and Moral Characters in Man, etc.* Biometrika, vol. 3, Part II, pp. 131-190.

Pearson, K...........'07 *On the Relationship of Intelligence to Size and Shape of Head, and to Other Physical and Mental Characters.* Biometrika, vol. 5, pp. 105-146.

Pedersen, R. H......'05 *Experimentelle Untersuchungen der visuellen und akustischen Erinnerungsbilder, angestellt an Schulkindern.* Archiv für die gesamte Psychologie, vol. 4, pp. 520-534.

Pillsbury, W. B......'03 *Attention Waves as a Means of Measuring Fatigue.* A. J. P., vol. 14, pp. 277-288.

Rice, J. M...........'97 *The Futility of the Spelling Grind.* The Forum, vol. 23, pp. 163-172 and 409-419.

Rice, J. M...........'02 *Educational Research: A Test in Arithmetic.* The Forum, vol. 34, pp. 281-297; *Causes of Success and Failure in Arithmetic.* The Forum, vol. 34, pp. 437-452.

Rice, J. M...........'03 *Educational Research, The Results of a Test in Language.* The Forum, vol. 35, pp. 269-293.

Rice, J. M...........'04 *English, The Need of a New Basis in Education.* The Forum, vol. 35, pp. 440-457.

Richardson, H........'02 *The Etiology of Arrested Mental Development.* Journal of Psycho-Asthenics, vol. 7, pp. 9-14.

Ritter, C.............'00 *Ermüdungsmessungen.* Zeitschrift für Psychologie, etc., vol. 24, pp. 401-444.

Rivers, W. H. R., and
Kraepelin, E..........'96 *Ueber Ermüdung und Erholung.* Psychologische Arbeiten, vol. 1, pp. 627-678.

Rosanoff, A. J., and
Orr, F. I.............'11 *A Study of Heredity in Insanity in the Light of the Mendelian Theory.* American Journal of Insanity, vol. 68, pp. 221-261.

Rusk, E. and H. M...'13 *Unpublished Study.*

Sakaki, Y............'05 *Ermüdungsmessungen in vier Japanischen Schulen.* International Magazine of School Hygiene, vol. 1, pp. 53-100.

Schuyler, W., and
Swift, E. J...........'07 *The Learning Process.* Psychological Bulletin, vol. 4, pp. 307-310.

Schuyten, M. C......'04 *Comment doit-on mesurer la fatigue des écoliers?* Archives de Psychologie, vol. 4, pp. 113-128.

Segal, J..............'08 *Uber den Reproduktionstypus und das Reproduzieren von Vorstellungen.* Archiv für die gesamte Psychologie, vol. 12, pp. 124-235.

Shipe, M. M......'03
Ellis, A. C., and *A Study of the Accuracy of the Present Methods of Testing Fatigue.* A. J. P., vol. 14, pp. 232-245.

Sikorski, J...........'79 *Sur les effets de la lassitude provoquée par les travaux intellectuels chez enfants le l'âge scolaire.* Annales d'hygiene publique, vol. 2, pp. 458-464.

Simpson, B. R.......'12 *Correlations of Mental Abilities.* Teachers College, Columbia University Contributions to Education, No. 53.

Spearman, C.....'04 (a) *The Proof and Measurement of Association between Two Things,* American Journal of Psychology, vol. 15, pp. 72-101.

Spearman, C.....'04 (b) *"General Intelligence" Objectively Determined and Measured.* American Journal of Psychology, vol. 15, pp. 201-292.

Spearman, C., and
Krueger, F...........'06 *Die Korrelation zwischen verschiedenen geistigen Leistungsfähigkeiten.* Zeitschrift für Psychologie, vol. 44, pp. 50-114.

Spearman, C., and
Hart, B..............'12 *General Ability, Its Existence and Nature.* British Journal of Psychology, vol. 5, pp. 51-84.

Specht, W............'04 *Ueber klinische Ermüdungsmessungen.* Archiv für gesamte Psychologie, vol. 3, pp. 245-340.

Starch, D............'11 *Transfer of Training in Arithmetical Operations.* Journal of Educational Psychology, vol. 2, pp. 306-310.

Stern, W............'00 *Uber Psychologie der individuellen Differenzen.*

Stern, W.............'11 *Die Differentielle Psychologie.*

Stone, C. W..........'08 *Arithmetical Abilities and Some Factors Determining Them.* Columbia University Contributions to Education, Teachers College Series, No. 10.

Swift, E. J., and Schuyler, W..........'07 *The Learning Process.* Psychological Bulletin, vol. 4, pp. 307-310.

Thompson, H. B......'03 *Psychological Norms in Men and Women.* University of Chicago Contributions to Philosophy, vol. 4, No. 1.

Thorndike, E. L......'00 *Mental Fatigue.* Psy. Rev., vol. 7, pp. 466-482 and 547-579.

Thorndike, E. L......'03 *Heredity, Correlation and Sex Differences in School Abilities.* Columbia University Contributions to Phil., Psy., and Ed., vol. 11, No. 2.

Thorndike, E. L......'04 *An Introduction to the Theory of Mental and Social Measurements.*

Thorndike, E. L......'05 *Measurements of Twins.* Archives of Philosophy, Psychology and Scientific Methods, No. 1.

Thorndike, E. L......'08 *Memory for Paired Associates.* Psychological Review, vol. 15, pp. 122-138.

Thorndike, E. L.....'08a *The Effect of Practice in the Case of a Purely Intellectual Function.* A. J. P., vol. 10, pp. 374-384.

Thorndike, E. L......'09 *The Relation of Accuracy in Sensory Discrimination to General Intelligence.* A. J. P., vol. 20, pp. 364-369.

Thorndike, E. L......'10 *Handwriting.* Teachers College Record, vol. 11, No. 2.

Thorndike, E. L......'10 *Practice in the Case of Addition.* A. J. P., vol. 21, pp. 483-486.

Thorndike, E. L......'11 *Mental Fatigue.* Journal of Educational Psychology, vol. 2, pp. 61-80.

Thorndike, E. L......'12 *The Curve of Work.* Psy. Rev., vol. 19, pp. 165-194.

Thorndike, E. L., and Donovan, M. E....'13 *Improvement in a Practice Experiment Under School Conditions.* A. J. P., vol. 24, pp. 426-428.

Thorndike, E. L., and Hahn, H. H..........'14 *Some Results of Practice in Addition under School Conditions.* Journal of Educational Psychology, vol. 5, No. 2.

Vannod, Th..........'96 *La fatigue intellectuelle et son influence sur la sensibilité cutanée.*

Vogt, R..............'99 *Ueber Ablenkbarkeit und Gewöhnungsfähigkeit.* Psychologische Arbeiten, vol. 3, pp. 62-201.

von Voss, G..........'99 *Ueber die Schwankungen der geistigen Arbeitsleistung.* Psychologische Arbeiten, vol. 2, pp. 399-449.

Wagner, L...........'98 *Unterricht und Ermüdung. Ermüdungsmessungen an Schülern des neuen Gymnasiums in Darmstadt.*

Ward, L. F..........'06 *Applied Sociology.*

Wells, F. L......'08 (a) *A Neglected Measure of Fatigue.* A. J. P., vol. 19, pp. 345-358.

Wells, F. L......'08 (b) *Normal Performance in the Tapping Test.* A. J. P., vol. 19, pp. 437-483.

Wells, F. L..........'12 *The Relation of Practice to Individual Differences.* A. J. P., vol. 23, pp. 75-88.

Wells, F. L..........'13 *Practise and the Work-Curve.* A. J. P., vol. 24, pp. 35-51.

Weygandt, W........'97 *Ueber den Einfluss des Arbeitswechsels auf fortlaufende geistige Arbeit.* Psychologische Arbeiten, vol. 2, pp. 118-202.

Whipple, G. M........'10 *Manual of Mental and Physical Tests.*

Whitley, M. T........'11 *An Empirical Study of Certain Tests for Individual Differences.* Archives of Psychology, No. 19.

Wiersma, H............. (See Heymans and Wiersma.)

Wimms, J. H........'07 *The Relative Effects of Fatigue and Practice Produced by Different Kinds of Mental Work.* British Journal of Psychology, vol. 2, pp. 153-195.

Winch, W. H........'10 *Some Measurements of Mental Fatigue in Adolescent Pupils in Evening Schools.* Journal of Educational Psychology, vol. 1, pp. 13-22 and 83-100.

Winch, W. H........'11 *Mental Fatigue in Day School Children as Measured by Arithmetical Reasoning.* British Journal of Psychology, vol. 4, pp. 315-341.

Winch, W. H........'12 *Mental Fatigue in Day School Children as Measured by Immediate Memory.* Journal of Educational Psychology, vol. 3, pp 18-28 and 75-82.

Wissler, C............'01 *The Correlation of Mental and Physical Tests.* Psychological Review, Monograph Supplement No. 16.

Wissler, C............'01 *The Growth of Boys.* American Anthropologist, New Series, vol. 5, pp. 81-88.

Woods, F. A..........'06 *Mental and Moral Heredity in Royalty.*

Woodworth, R. S....'10 *Racial Differences in Mental Traits.* Science, New Series, vol. 31, pp. 171-186.

Yoakum, C. S........'09 *An Experimental Study of Fatigue.* Psychological Review, Monograph Supplement No. 46.

INDEX