EDUCATIONAL PSYCHOLOGY

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO

MACMILLAN & CO., LIMITED
LONDON · BOMBAY · CALCUTTA
MELBOURNE

THE MACMILLAN CO. OF CANADA, LTD.
TORONTO

EDUCATIONAL PSYCHOLOGY

BY

DANIEL STARCH, PH. D.

UNIVERSITY OF WISCONSIN

New York

THE MACMILLAN COMPA

1924

PRINTED IN THE UNITED STATES OF AMERICA

Set up and electrotyped. Published June, 1919. Reprinted January, April, August, December, 1920; November, 1921; October, 1922; April, October 1923; January, 1924.

NTING COMPANY

ORK

TO MY WIFE

PREFACE

The preparation of this book has been carried out according to two fundamental purposes: First, to present that material which seems to be most useful and relevant to the problems of educational psychology; and second, to maintain a strictly experimental, scientific viewpoint in discussing these problems. The result of these aims has been a considerable reduction in the amount of space usually devoted in texts on educational psychology to certain topics such as, instinct, fatigue, and imagery, and the inclusion of new topics such as tests of intelligence, studying, transference of training in school subjects, the assignment of marks, and much of the material in Part III which has as yet not found a place in text-books.

The space devoted to the discussion of instinct has been materially reduced for two reasons: In the first place, while the instincts are fundamental in human life, too much time has usually been devoted to their consideration for the amount of direct benefit gained. The actual use in school work that can be made of a detailed knowledge of instincts, which in our present stage of information is largely analytical and theoretical, is relatively small when it comes to dealing face to face with concrete school problems. In the second place, a great deal of experimental and statistical material has accumulated in recent years which is more immediately valuable in solving the problems of the psychology and pedagogy of learning.

It would have been desirable to include a discussion of the psychology of more of the high school subjects; but this is impossible at the present time. The discussion of the school subjects in Part III has been confined to tangible, scientific investigations. Obviously there is little or no material of this sort on most of the high school subjects. The consideration of educational tests in the chapters of Part III is perhaps brief; but a detailed treatment of the theoretical and statistical principles underlying their construction belongs rather in special treatises. Chapter XII on How to Study is not altogether satisfactory, because of the scarcity of definite or substantial material in this field. It was, however, included because the topic is exceedingly important in school work and because it was hoped that its inclusion would stimulate discussion of it by teachers and prospective teachers.

I take pleasure in expressing my obligations to the persons who have assisted me in various ways in the preparation of this book; namely, to Dr. Helen Hubbert Caldwell and Mr. A. O. Hansen, who have read the manuscript and offered many helpful suggestions, to Mr. W. R. Ames who has prepared the drawings, and especially to Dr. C. L. Hull who has critically examined every portion of the manuscript and has offered many suggestions which have been incorporated in the book.

DANIEL STARCH.

MADISON, WISCONSIN,
October 5, 1918.

CONTENTS

EDUCATIONAL PSYCHOLOGY

EDUCATIONAL PSYCHOLOGY

CHAPTER I

PROBLEMS AND SCOPE OF EDUCATIONAL PSYCHOLOGY

What is Education? The problems and the scope of educational psychology are necessarily determined by our conception of what education is. If we conceive education to be primarily self-development, our problems will be of one sort; if we conceive education to be fundamentally social adaptation, our problems will be of another sort. In the former case, education would mean the complete training of the mental and physical capacities irrespective of environment; in the latter case, education would mean the training of those capacities which will adapt the individual most adequately to the social and physical environment in which he is to live. For our present purpose it is not necessary to define in complete detail the aim and meaning of education. It will be sufficient to state in the simplest terms what education is as a psychological process.

In the broadest sense, education is the production of useful changes in human beings.[1] These changes may be classified into three divisions: changes in knowledge, in skill, and in ideals. Through education the child is to acquire useful knowledge; he is to acquire skill, both motor and intellectual, in the use of his muscles and in the manipulation of ideas and concepts; and, finally, he is to acquire the right ideals of life which will actually function in his behavior. Probably all changes wrought in human beings which in any sense are educational, fall under these three heads. Obviously then, education is the most momentous, as well as the most essential, business of the human race; for the welfare of the race depends upon education as it depends upon nothing else.

[1] Thorndike has defined the purpose of education thus: "The aim of education is, as we have seen, to change human beings for the better, so that they will have more humane and useful wants and be more able to satisfy them." ('12, p. 52.)

Which changes are useful and which are not is a question that cannot be answered as easily as it would seem at first glance. Learning to read or to figure are obviously useful modifications; but it is not so easy to say whether the study of a given drama, or the knowledge of certain facts of history, or the understanding of a certain theory of matter, or ability to read a given foreign language, are useful, or sufficiently useful to be included in the common school, in the high school, or in the university. The term useful should not be limited narrowly to the things which are immediately applicable in making a living, but should include all changes which will broaden and enrich the life of the individual.

The Problems of Educational Psychology. In accordance with our definition, the fundamental problems that we must raise concerning education are as follows:

1. What changes are to be made in human beings?
2. What are the agencies by which the changes may be brought about?
3. What are the capacities which human beings possess for acquiring the changes?
4. What are the most economical methods by which these changes may be brought about?

The first problem is primarily a problem for philosophy and sociology. What changes are ultimately to be made depends upon our ideals of life and our views of society. The modifications that have been sought by different nations and different systems of education have varied from century to century and from race to race. The ultimate aims of education sought by the ancient Greeks or by the mediæval monks were very different from those sought by the modern Americans or Europeans.

The second problem deals primarily with the value of school subjects and exercises in bringing about the changes that are to be made. To what extent will the study of arithmetic, the study of grammar, or the study of physics or Latin be able to produce the training that philosophy and sociology dictate? This problem is partly sociological and partly psychological. It is sociological in so far as the determination of educational agencies depends upon the physical and social environment of mankind; it is psychological in so far as it necessitates a study of mental processes affected or brought about by these agencies. This latter phase of the problem merges into problems three and four.

Problems three and four are fundamentally psychological and, together with the second phase of problem two, constitute the main scope of educational psychology. It is a psychological problem to determine what capacity and equipment human beings have for acquiring the changes that are to be made. Likewise, it is fundamentally a psychological problem to discover the most economical methods of learning. Accordingly, the field of educational psychology is divided into two large divisions which we may designate as:

I. The native equipment of human beings;
II. The psychology of learning.

Psychology and Teaching. Methods of teaching rest fundamentally upon the psychology of learning. Since the experimental analysis of learning processes will have to reveal the principles according to which the human mind learns, and learns most economically, it follows that the methods of teaching will have to be based upon these discoveries. This may be illustrated in the case of reading. If the child learns to read most economically by the word method, it follows that the most economical way of teaching reading would be by the word method. Likewise, if a child learns to spell homonyms more easily by studying them together, or memorizes prose or poetry more readily by wholes than by parts, it follows that these exercises should be taught accordingly. Evidently the fundamental principles of teaching must be based upon the psychological laws and principles of economic learning.

Waste in Education. Exact information concerning the proper procedure in educational matters is exceedingly rare. Definite, scientific knowledge of the proper methods of learning and teaching school subjects and of the efficient administration of our schools is surprisingly small, and the field of educational psychology in its broadest sense opens up endless problems for the future to solve. We know relatively little in a scientific way about the learning of any single school subject. For example, we do not know with any definite assurance what is the most economical amount of time to devote to any one of the school subjects. From such investigations as have been made, we may infer that there is an enormous waste in our educational practices which is indicated by such facts as the following: It has been found by recent tests and measurements that some schools obtain just as good results by devoting only one-half as much time to writing as other schools do. Similar facts have been brought out in the case of reading, arithmetic and other school subjects. Schools which have devoted as much as 100 minutes a week, or 20 minutes a day, have obtained no better results than other schools devoting 50 minutes a week, or 10 minutes a day, to the same subject. If these facts actually represent the real possibilities, it seems quite obvious that there is an enormous waste in our schools and this waste is far greater than we realize until we make definite calculations of the possible saving of time. If by some means it were possible to save one minute a day for every school day during the eight years of a child's school life, we would be able to save one entire week of school time. If we could save four minutes a day for the same length of time, we would be able to save one month; if we were able to save 18 minutes a day, we would be able to save one-half of a school year; and if by more economical methods of learning and distribution of time we were able to save 36 minutes a day for eight years, we would be able to save an entire school year. Such a saving is not impossible; indeed, by a better use of time and more effective methods of learning, it is highly probable. Eighteen minutes a day would mean a reduction of only 4½ minutes in each of four subjects; 36 minutes a day would mean a reduction of only 9 minutes a day in each of four subjects. This time could be devoted with greater advantage to other and possibly more advanced school subjects and school exercises.

The Specific Topics and Problems. In order that we may be properly orientated with reference to the problems that will be discussed under the two large divisions of educational psychology, the following order of topics will be considered:

I. The native equipment of human beings.
   - a. What does it consist of?
   - b. To what extent does it vary?
     - 1. Among individuals.
       - (a) In single traits.
       - (b) In combinations and relationships of traits.
     - 2. At different times of life in the same individual.
     - 3. Between the sexes.
   - c. To what extent is it inherited?
   - d. How may it be measured?

II. The Psychology of Learning.
- a. The psychology of learning in general.
  1. Observation and perception.
  2. The rate and progress of learning.
  3. Transference of training.
- b. The psychology of learning school subjects in particular.
  1. The psychological processes involved in each subject.
  2. The measurement of ability and progress in learning each subject.
  3. The most economical methods of learning the material of each subject.

PART I

THE NATIVE EQUIPMENT OF HUMAN BEINGS

CHAPTER II

THE INSTINCTIVE ELEMENTS OF NATIVE EQUIPMENT

Reflexes, Instincts, and Capacities. The equipment with which human beings start in life may be divided into three types of inherited responses and abilities: Reflexes, instincts, and capacities. The distinction among these three is primarily one of definiteness and degree of complexity. An instinct may be defined from the neurological side as an inborn neural connection between sense organ and muscle. It may be defined from the functional side as an inborn capacity of responding in definite ways under definite circumstances. These responses are prior to experience and training, and need not be learned. To close the eyes when an object suddenly approaches them, to get food when hungry, to strike when struck, and to be afraid of thunder and of the dark, are illustrations of instinctive responses. The reflexive and instinctive responses are inherited in the sense that there is present in the nervous system, either at the time of birth or later on as a result of growth, a set of nervous connections already formed for the carrying out of a particular action in response to a given situation. If the child closes his eyes when an object suddenly approaches, it means that the motor impulses travel from the retina to the visual center of the brain, from there to the motor center which controls the movement of the eyelids, and from there out to the muscles of the eyelids to cause the contraction. In the case of inherited responses, the connection from the sensory to the motor centers is already present and ready to operate in carrying out the action. In the case of acquired responses, such as habits, these nervous connections must be formed as a result of effort and trial on the part of the individual.

The difference between reflexes and instincts is largely a difference of complexity. Both are inherited types of responses. Reflexes are simpler forms of reaction usually involving a limited set of muscles and occurring in response to precise stimuli. The contraction or expansion of the iris, the closing of the eyelids, the knee jerk, are illustrations of reflexes. Instincts are complex reactions involving the use of large groups of muscles or, in many instances, the entire muscular system of the body. They may be aroused either by external stimuli or situations or possibly by internal stimulation. To make movements in the direction of getting food when hungry, to seek shelter when cold, to offer resistance when hemmed in, to spit out what tastes bad, and the like, are instinctive reactions. Capacities are distinguished from reflexes and instincts in being general mental abilities rather than specific motor responses and in referring primarily to the native mental equipment, such as the powers of sensation, perception, retention, attention, imagination, and all the varied complex psychic processes.

Classification of Instinctive Responses. The older classifications of instincts usually divided them into three or four large groups of responses, such as individual, racial, and social, and regarded them rather as general tendencies than as specific responses. The present conception of instincts is to regard them as specific responses with inherited neural mechanisms which will be set into action by specific stimuli or situations. On this basis the classification consists of an enumeration of as many definite, identifiable, unlearned reactions to specific situations as can be observed and as can be recognized in human beings prior to training and habituation in each particular type of activity. Accordingly, Thorndike ('14, I) enumerates forty or more different types of instinctive reactions as follows:

I. Food getting and protective responses.
   1. Eating.
   2. Reaching, grasping, and putting objects into the mouth.
   3. Acquisition and possession.
   4. Hunting.
   5. Collecting and hoarding.
   6. Avoidance and repulsion.
   7. Rivalry and coöperation.
   8. Habitation.
   9. Response to confinement.
   10. Migration and domesticity.
   11. Fear.
   12. Fighting.
   13. Anger.

II. Responses to behavior of other human beings.
14. Motherly behavior.
15. Gregariousness.
16. Responses of attention to other human beings.
17. Attention getting.
18. Responses to approving and to scornful behavior.
19. Responses by approving and scornful behavior.
20. Mastering and submissive behavior.
21. Display.
22. Shyness.
23. Self-conscious behavior.
24. Sex behavior.
25. Secretiveness.
26. Rivalry.
27. Coöperation.
28. Suggestibility and opposition.
29. Envious and jealous behavior.
30. Greed.
31. Ownership.
32. Kindliness.
33. Teasing, tormenting, and bullying.
34. Imitation.

III. Minor bodily movements and cerebral connections.
35. Vocalizations.
36. Visual exploration.
37. Manipulation.
38. Cleanliness.
39. Curiosity.
40. Multiform mental activities.
41. Multiform physical activities.
42. Play.

Relation of Education to Native Endowment. The inherited equipment of the human being is the foundation upon which education must build; it consists of the faculties and capacities which the child has for reacting to his environment. It is the utilization, the training or the curbing of these endowments which education attempts to accomplish. In much of the writing and thinking concerning educational problems, there has been a relative overemphasis, in space and time, upon instincts and an underemphasis upon the mental capacities. Education in the sense of schooling has as much if not more to do with the latter than with the former. The direct appeal to, and use of, instinctive reactions in actual concrete instances in school work are not as frequent and specific as is commonly implied. The number of instincts enumerated in the preceding list which may be directly and concretely appealed to in the learning of a school subject is relatively small. The best way to be convinced on this point is to take the various instincts one by one, and to determine to what extent each one may be appealed to or used in teaching the various subjects. The number of specific applications is much smaller than one is likely to anticipate. Two-thirds or three-fourths of them are probably never immediately but only indirectly concerned in school exercises, and most of the remaining ones, such as rivalry, coöperation, collecting and hoarding, are serviceable chiefly as general motives. As such they are, to be sure, highly important.

We must, of course, not minimize the place and importance of instinctive reactions in behavior as a whole. They furnish the general motives and mechanisms for doing and learning, but the mental capacities are more directly and concretely involved in the acquisition of knowledge and skill in school subjects.

The instinctive elements in learning any school subject are for the most part simple reflex actions or undeveloped connections. Take learning to read as an illustration. The chief instinctive elements probably are the reflexes in the control of the eyes, the neural mechanism for receiving and transmitting visual impulses to the brain, the capacities for attentiveness and retentiveness, and partial motor control of the speech organs. The process of learning to read assumes these and uses them; but, what is more important from the practical side of getting the meaning of the printed word is the establishment of countless new connections.

Perhaps the most important rôle of instinct in education lies in motivating and energizing the learning processes. There can be no education except through the activity of the child himself; and no activity can take place which does not ultimately depend upon native tendencies. They are the origin of effort, the springs of action. The skillful teacher plays upon them and appeals to them in countless ways. The ability to do this is an art which is not easily learned from books; it is acquired rather by patient practice and by sympathetic contact with children.

The energizing power of instinct makes itself felt largely through its control of the attention processes. Owing to the peculiarities of our inherited nervous organization, certain impressions have a potency over others in attracting the attention and interest of the child. A flash of lightning, a holiday parade, one's name in the newspaper, or a moving picture make certain instinctive appeals to the attention of a young girl which the study-lamp, the doing of errands for mother, the seeing of a stranger's name in the newspaper, or the reading of the history lesson do not make. The great importance of attention for the learning process lies in the fact that associations, analyses, and indeed all mental processes are carried out much more effectively when they occupy the focus of attention. Ebbinghaus found that, after inattentively reading over nonsense syllables until many successive persons had learned them perfectly, he himself could repeat very few of them. Impressions must occupy the focus of consciousness in order to be retained effectively.

Considering the three main types of attention, passive, active, and secondary passive, the most simple and the one most directly related to the instinctive life is probably the first. Passive attention is such as one gives spontaneously to any curious or interesting sight or sound. Active attention is such as one gives perhaps to an inherently distasteful task which requires an effort of will to keep the mind upon it. While such a task itself does not supply the stimulus for vigorous instinctive reaction, it is in some direct way connected with one that does. A little girl will apply herself to the disagreeable task of learning a spelling lesson, not because the words in themselves have any charm for her but because she has the instinctive craving for the approval of her teacher. The third type of attention, secondary passive or derived attention, is attention which has become passive only after having passed through an initial active stage. It is illustrated by the common experience of becoming absorbed in a task which at first required a distinct effort. In the beginning the motivation lay outside the task, say in a sense of duty or social obligation; but after a time an adequate stimulus for activity was encountered in the work itself.

Apparently back of every act of attention lies somewhere a more or less primitive, innate tendency to action. To focus the attention of a class upon various associations involved in learning the multiplication table, a skillful teacher may on one day appeal to curiosity in the novelty of the combinations; on another day she may appeal to the native pleasure in rhythm by making the table into a rhyme or song. Or the tendency to play will be utilized by making the number combinations into a game. More than likely the game itself will depend upon instinctive rivalry and emulation. The love of social approval is appealed to by giving distinctions, marks, gilt stars, and the like. Future advantage may be used as an inducement for present application. The instinct of pugnacity may be utilized in wanting to succeed in a hard task. The teacher herself, in standing position with face and body in animated attitudes, may appeal to the fundamental interest in change and action. Lastly may be mentioned the more doubtful negative motives of deprivation from coveted privileges and, biologically perhaps the most fundamental and powerful of all motives, physical pain. It may be worth noting that animal psychologists have found pain in some cases a more potent motive to learning than pleasure. Hoge and Stocking ('12) found that rats when rewarded by food alone had by no means learned perfectly certain sensory habits in 610 trials; when they were punished for failures they learned the habits perfectly in this number of trials, but when they were both rewarded for successes and punished for failures they learned the habits perfectly in 530 trials.

Educational Doctrines Based upon Instincts. Some very far-reaching speculations and theories with regard to the nature and value of instincts for education have been spun out, some of which are largely imaginary and questionable, and are based upon analogy rather than fact. For the most part these educational doctrines have centered around three concepts.

The first is that instincts are the great dynamic forces of human nature which determine the actions, desires, and achievements in an individual's life. Hence the injunction to the school has been to work with nature rather than against her or apart from her. We shall call this the dynamic theory of instincts.

The second is that these instincts are highly transitory; that they burst out at certain times in the growth of the individual with more or less dramatic force and suddenness, and that if they are not allowed to manifest themselves, they will disappear never to be revived again. From this assumption has been derived the pedagogical application of the maxim, "strike while the iron is hot." This is the theory of the transitoriness of instincts.

The third is that instincts appear in the growth of the child in the order in which they appeared in the evolution of the race. From this assumption has been derived the pedagogical maxim, "teach the child his activities in the order in which the interests for them appear." This is known as the recapitulation theory of instincts.

Critique of the Dynamic Theory of Instincts. To work with nature rather than against her is undoubtedly a sound principle. The fundamental instincts of man are the driving forces of human life that determine ultimately the motives and causes of behavior. They are so deep-seated in the human psychophysical organism that we may almost say that to work apart from, or against, nature is a futile task. If, through the instinct of multiform activities, a child manifests a tendency to draw, the school should take advantage of this original impulse and build upon it. All the original manifestations of a child's nature should be used in the acquisition and training of those exercises which education considers valuable.

This dynamic theory of instincts, however, involves on the one hand a difficulty and on the other a questionable assumption. The difficulty is that the principle is general and as such does not point out the particular ways in which the school may coöperate with the inborn forces of child nature. It is easy enough to say "work with nature," but just how is that to be done in teaching a pupil how to make the letter "a," or to learn the reading of a printed word, or to acquire correct speech, or to learn the grammatical rules of a foreign language? Ultimately, the concrete use of the principle must be determined experimentally. Our definite knowledge of the technique of learning in the case of school subjects is appallingly limited. Only by careful and painstaking experimentation can this principle be made useful in the concrete work of the school in anything more than an offhand impressionistic manner.

The questionable assumption is that the instincts are infallible guides of human life. It may be argued that since instincts are such powerful springs to action as to have maintained the individual and the race for numberless generations, they must necessarily be dependable in producing action and interest of the right sort. But the question may fairly be raised: Are the native tendencies always right so that we should always coöperate with them and never counteract or curb them? The theory of the infallibility of instincts is based on the belief that for countless ages nature has found by experimentation and natural selection what is best for the individual. Whatever the child is inclined to do by virtue of his natural proclivities is right and good for him; or, if apparently not useful, it is a necessary precursor or a necessary accompaniment of useful tendencies. On the theory of immunization or catharsis, the undesirable tendencies prepare the ground for the proper development and growth of the desirable ones. However, the theory of catharsis is highly questionable and runs counter to the law of use and disuse which in general operates to the effect of making permanent the functions exercised. The belief is that if a boy has proclivities toward thieving or lying or dishonesty, and is allowed to exercise these unconstrained, he will purge himself of these tendencies and be the more honest and truthful later on in life. Concrete data, however, in addition to the general principle of use and disuse, seem to point in the opposite direction. Experimental and statistical inquiries show that the relative strength of the various traits remains fairly constant throughout life; that if a child manifests certain abilities and tendencies even during childhood, these abilities and tendencies will remain relatively dominant during his adult life. Early interests and intellectual capacities are very certain, on the whole, of being prophecies of similar interests and intellectual abilities later in life. As will be pointed out in a succeeding chapter, scholastic ability remains fairly constant in any given child all through his educational career.

Furthermore, we must remember that nature, in the development of instincts and in securing adaptation to environment, works on the whole in a very slow and prodigal manner; and that conditions of life may change long before the organism through its behavior appropriately adapts itself to the surroundings by furnishing the necessary native equipment within the organism. On this account a great deal of our native equipment is out of date and has adapted man to primitive conditions of uncivilized life. As a result of this, we manifest many tendencies which are not particularly useful at the present time. We would be better off if in place of them we had instinctive capacities for meeting situations with which we are to-day confronted in civilized life. As Thorndike has pointed out:

"The imperfections and misleadings of original nature are in fact many and momentous. The common good requires that each child learn countless new lessons and unlearn a large fraction of his natural birthright. The main reason for this is that original equipment is archaic, adapting the human animal for the life that might be led by a family group of wild men in the woods, amongst the brute forces of land, water, wind, rain, plants, animals, and other groups of wild men. The life to which original nature adapts man is probably far more like the life of the wolf or ape, than like the life that now is, as a result of human art, habit and reasoning, perpetuating themselves in language, tools, buildings, books and customs." ('14, I, p. 280.)

That these primitive tendencies persist with great strength is shown by the ready manner in which the veneer of civilization comes off and by the fact that men and women in strained circumstances will easily revert to their primitive, brutal instincts. The chief advocate of the theory of infallibility and catharsis of instincts has been G. Stanley Hall, who has said:

"In education, don't cut off the tadpole's tail."

"Rousseau would leave prepubescent years to nature and to these primal hereditary impulsions and allow the fundamental traits of savagery their fling till twelve. Biological psychology finds many and cogent reasons to confirm this view *if only a proper environment could be provided.* The child revels in savagery; and if its tribal, predatory, hunting, fishing, fighting, roving, idle, playing proclivities could be indulged in the country and under conditions that now, alas! seem hopelessly ideal, they could conceivably be so organized and directed as to be far more truly humanistic and liberal than all that the best modern school can provide. Rudimentary organs of the soul, now suppressed, perverted, or delayed, to crop out in menacing forms later, would be developed in their season so that we should be immune to them in maturer years, on the principle of the Aristotelian catharsis for which I have tried to suggest a far broader application than the Stagirite could see in his day." (Hall, '08, p. 2.)

"He should have fought, whipped and been whipped, used language offensive to the prude and to the prim precisian, been in some scrapes, had something to do with bad, if more with good, associates, and been exposed to and already recovering from as many forms of ethical mumps and measles as, by having in mild form now he can be rendered immune to later when they become far more dangerous, because his moral and religious as well as his rational nature is normally rudimentary." ('08, p. 235.)

The Theory of the Transitoriness of Instincts. This conception of instincts contains two elements: First, the suddenness of the appearance of instincts, and second, the unrevivable disappearance of instincts. The most conspicuous advocate of the former idea in this country, has been G. Stanley Hall, who states it as follows:

"But with the teens all this begins to be changed and many of these precepts must be gradually reversed. There is an outburst of growth that needs a large part of the total kinetic energy of the body. There is a new interest in adults, a passion to be treated like one's elders, to

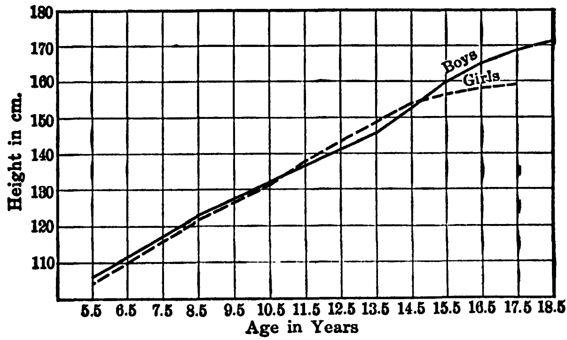

FIG. 1.—Height of boys and girls measured in centimeters, based on measurements of 45,151 boys and 43,298 girls. After Boas ('96–'97).

make plans for the future, a new sensitiveness to adult praise or blame. The large muscles have their innings and there is a new clumsiness of body and mind. The blood-vessels expand and blushing is increased, new sensations and feelings arise, the imagination blossoms, love of

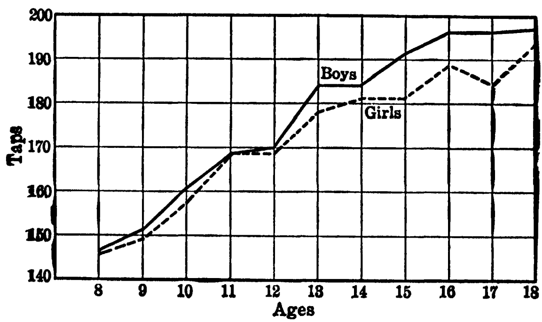

FIG. 2.—Rate of tapping. Number of taps made with right hand in 30 seconds. Based on tests with 395 boys and 495 girls. After Smedley ('00–'01).

nature is born, music is felt in a new, more inward way, fatigue comes easier and sooner; and if heredity and environment enable the individual to cross this bridge successfully, there is sometimes almost a break of continuity, and a new being emerges." ('08, p. 236.)

The correctness of the theory of sudden appearance is primarily a question of fact. Thus Hall describes the social instincts at the time of adolescence as follows:

"The social instincts undergo sudden unfoldment and the new life of love awakens. It is the age of sentiment and of religion, of rapid fluctuation of mood, and the world seems strange and new. Interest in adult life and in vocations develops. Youth awakes to a new world and understands neither it nor himself." ('04, Preface p. XV.)

The advocates of this viewpoint maintain, therefore, that there is a nascent period for motor activity, for memory and habituation, for reason and logical thinking and the like; that the school should seize these opportunities to teach those activities which will exercise the particular capacities that occupy the stage of youth at that period; and that more can be accomplished at those periods in a given length of time than can be accomplished in several fold as much time later on.

The facts do not seem to warrant an interpretation of such marked suddenness but indicate rather a gradual waxing of instincts. There appears to be no special time during which the child suddenly begins to reason or to reason very much more than he had done theretofore. The same description seems to be true of memory, motor ability, the collecting instinct, and other capacities, as indicated in the accompanying graphs.

A great deal of the dramatic bursting forth of instincts is chiefly a dramatic bursting forth of descriptive words. The actual facts seem to justify more nearly an interpretation of gradual unfoldment instead of a sudden bursting forth. Growth in height and weight proceeds by a very uniform increase even during the adolescent period. Motor capacity grows steadily and uniformly without particularly sudden leaps or bounds. Memory ability increases steadily for both rote and logical material up to adulthood, during which it probably remains fairly constant until senility sets in. There is no memory period during which the child memorizes very much more readily than he did before or than he ever will later. The memory of the average adult for both mechanical and logical material and for either immediate or permanent retention is superior to the memory of the average child at any age. Even reasoning ability, which is usually described as appearing suddenly at the dawn of adolescence, is a matter of gradual development. To argue

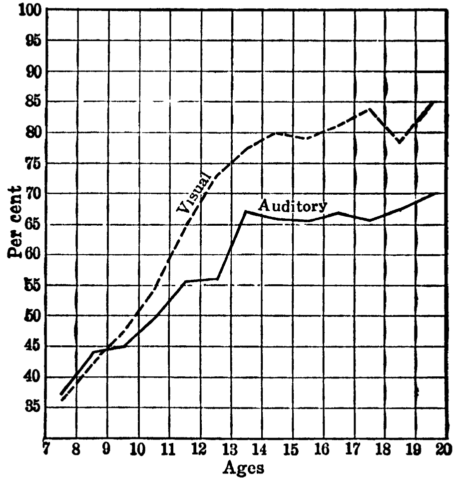

FIG. 3.—Memory for digits based upon tests of 937 pupils. After Smedley ('00–'01).

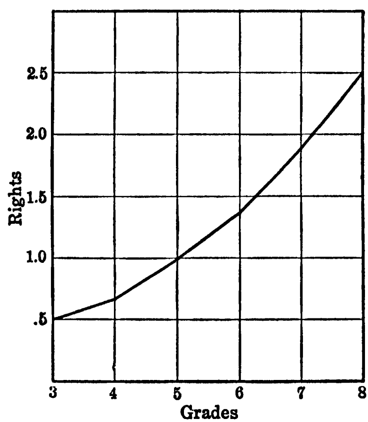

FIG. 4.—Development in arithmetical reasoning as measured by the Courtis test No. 8, Series A. The vertical axis shows the number of problems done correctly in six minutes.

that school exercises which consist mainly of memorizing should be placed at the "memory age" on the ground that the pupil will learn them more readily at that time than at a later time in life, is fallacious. It may be advisable to begin the study of foreign languages earlier than is customary, but not for any reason of more rapid memorizing at an earlier age. If rapidity of tapping, Figure 2, is any indication at all of endurance or of quickness of becoming fatigued, it does not seem to be true that "fatigue comes easier and sooner" during the adolescent stage. The graphs for tapping

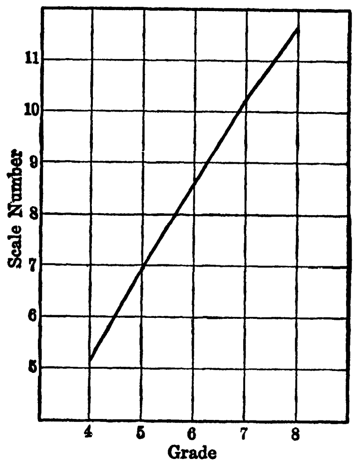

FIG. 5.—Development in arithmetical reasoning as measured by the arithmetical scale A. After Starch ('16). The vertical axis represents the scale step or the number of problems done correctly.

do not drop but tend to rise gradually even during the years from eleven to fifteen. There is practically a level at the age of eleven but no drop.

The unrevivability of instincts through disuse has been advocated chiefly by James as follows:

"This leads us to the law of transitoriness which is this: Many instincts ripen at a certain age and then fade away. A consequence of this law is that if, during the time of such an instinct's vivacity, objects adequate to arouse it are met with, a habit of acting on them is formed, which remains when the original instinct has passed away; but that if no such objects are met with, then no habit will be formed; and, later on in life, when the animal meets the objects, he will altogether fail to react, as at the earlier epoch he would instinctively have done." ('90, II, p. 398.)

"Leaving lower animals aside, and turning to human instincts, we see the law of transiency corroborated on the widest scale by the alteration of different interests and passions as human life goes on. With the child, life is all play and fairy-tales and learning the external properties of 'things'; with the youth, it is bodily exercises of a more systematic sort, novels of the real world, boon-fellowship and song, friendship and love, nature, travel and adventure, science and philosophy; with the

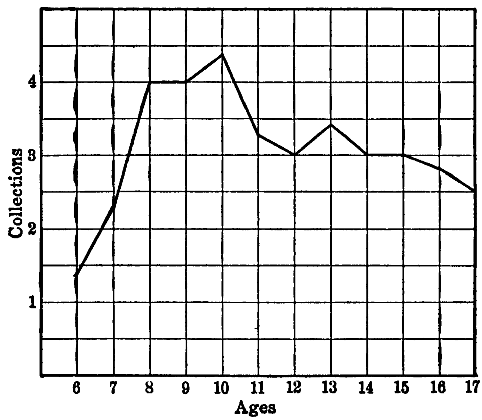

FIG. 6.—The number of collections made by children at various ages. After C. F. Burk ('00).

man, ambition and policy, acquisitiveness, responsibility to others, and the selfish zest of the battle of life. If a boy grows up alone at the age of games and sports, and learns neither to play ball, nor row, nor sail, nor ride, nor skate, nor fish, nor shoot, probably he will be sedentary to the end of his days; and, though the best of opportunities be afforded him for learning these things later, it is a hundred to one but he will pass them by and shrink back from the effort of taking those necessary first steps the prospect of which, at an earlier age, would have filled him with eager delight. The sexual passion expires after a protracted reign; but it is well known that its peculiar manifestations in a given individual depend almost entirely on the habits he may form during the early period of its activity. Exposure to bad company then makes him a loose liver all his days; chastity kept at first makes the same easy later on. In all pedagogy the great thing is to strike the iron while hot, and to seize the wave of the pupil's interest in each successive subject before its ebb has come, so that knowledge may be got and a habit of skill acquired—a headway of interest, in short secured, on which afterward the individual may float. There is a happy moment for fixing skill in drawing, for making boys collectors in natural history, and presently dissectors and botanists; then for initiating them into the harmonics of mechanics and the wonders of physical and chemical law. Later, introspective psychology and the metaphysical and religious mysteries take their turn; and, last of all, the drama of human affairs and worldly wisdom in the widest sense of the term." ('90, II, pp. 400–401.)

Facts seem to support this aspect of the nature of instincts more than the theory of the sudden appearance. The law of disuse of functions is necessarily in general support of the theory. Any functions will, as a rule, be strengthened through exercise. The assumption, however, that instincts, if not exercised when they first manifest themselves, will become dormant beyond the possibility of reawakening, or that they actually become dormant, is questionable. James gives isolated instances in favor of his viewpoint. Experimental and comprehensive observations are missing at the present time. Isolated illustrations of the opposite viewpoint, however, also are to be found. Thus it frequently happens that through the change of circumstances of life, instincts apparently long dormant or never given opportunity to manifest themselves, will quickly appear for action. For example, the writer has a friend who as a boy had never developed the instinctive tendencies involved in fishing. About the age of thirty, through the opportunities of a new environment, the instinct appeared so strongly that he will go to great lengths at any time of day or night to follow this sport. But isolated instances are dangerous bases on which to generalize, and future inquiries will have to solve the problem. Many instincts apparently are dormant only because no opportunity of expressing themselves are at hand, or because other more dominant interests prevail, but may, when appropriate circumstances arise, rapidly appear for action.

The Recapitulation Theory of Instincts. The principle of recapitulation was formulated by biologists to account for the development of animal organisms in the early stages of their growth. The theory of recapitulation holds that the individual retraces in its growth the successive stages of development of the entire animal series. Thus Hall says:

"Holding that the child and the race are each keys to the other, I have constantly suggested phyletic explanations. . . ." ('04, I, p. VIII.)

"The best index and guide to the stated activities of adults in past ages is found in the instinctive, untaught, and non-imitative plays of children. . . . In play every mood and movement is instinct with heredity. Thus we rehearse the activities of our ancestors, back we know not how far, and repeat their life work in summative and adumbrated ways. It is reminiscent, albeit unconsciously of our line of descent, and each is the key to the other. . . . Thus stage by stage we enact their (our ancestors') lives. Once in the phylon many of these activities were elaborated in the life and death struggle for existence. Now the elements and combinations oldest in the muscle history of the race are re-presented earliest in the individual, and those later follow in order." ('04, I, pp. 202–203.)

And Puffer says: "We are by turns vertebrates, gill-breathing vertebrates, lung-breathing vertebrates (we make the great change at birth), little monkeys, little savages, and finally civilized men and women." ('12, p. 77.)

The evidences given for the principle of recapitulation are largely embryological and structural. Vestigial organs such as the vermiform appendix, gill slits, etc., are further cited as evidences of the remainder of structures once useful. The facts seem to be that such recapitulation as takes place is very brief and confined almost wholly to the prenatal period of an individual's development. Davidson, after a comprehensive review of the biological evidence for the theory, concludes thus:

"The history of recapitulation is an instructive one. A principle of limited application within the field of its origin was elevated to a position of wide generality, and so gave rise to a conception in the main misleading. Carried into a new territory without a sufficient examination of its merits, it was applied broadly as an explanatory principle and thus distributed its misleading influence beyond its own borders." ('14, p. 99.)

"A more thorough consideration of the facts has led to a view of development essentially contradictory to this recapitulatory one. Ontogeny represents the ancient life-cycle which as such has been transmitted from the beginning. The chronological sequence from egg to maturity is not a rehearsal of a like historical series of events throughout the phylogeny of species; it is but the recurrence of an order which has been repeated in the lifetime of each individual from the beginning. In general, the effect of the modifications induced by germinal mutations and selection in the successive ontogenies, to make them over, and to destroy the resemblance of later ones to their predecessors."

The recapitulation theory with its pedagogical corollate, the culture epochs theory, has been developed largely as an analogy with many of the analogues missing. Its usefulness for educational thinking seems to the writer to be greatly exaggerated. It has built pedagogical mountains out of biological molehills. It is primarily an anatomical principle proposed to account for the embryological development of biological organisms, and has been brought over into human behavior to explain on the one hand, the order and dates of appearances of instincts, and to furnish on the other hand, a basis for the order and dates of teaching subjects in the school curriculum. The former assumption is more or less dubious, since most, if not all, of the demonstrable recapitulation occurs before birth; and the latter assumption is quite certainly dubious, since the anatomical and probably also the functional recapitulation has long ceased when the child begins his definite education.

CHAPTER III

VARIATION IN HUMAN CAPACITIES

How May They be Measured and Represented? Differences among human beings are quantitative rather than qualitative. That is, all human beings have the same reflexes, instincts, and capacities; all have the powers of perception, discrimination, attentiveness, retentiveness, reasoning, and so forth. All persons, consequently, have the same general qualitative make-up. The variations from person to person are, therefore, primarily differences in the strength of the various abilities that each individual possesses, and in the manner in which amounts of the various

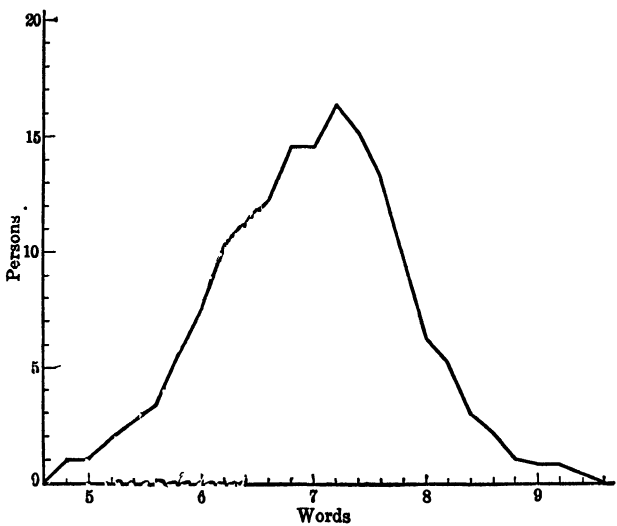

FIG. 7.—Distribution of memory ability of 173 University students. The test consisted in dictating ten monosyllabic nouns. The persons then recorded the words that they remembered. The horizontal axis indicates the number of words and the vertical axis indicates the number of persons having each memory ability.

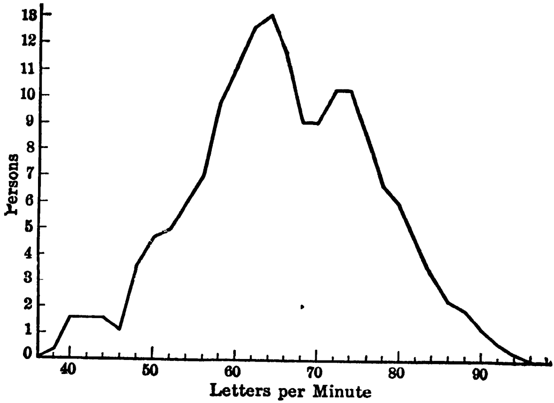

FIG. 8.—Distribution of ability in the A-test. Based on 164 University students. The horizontal axis represents the number of A's canceled in one minute; the vertical axis represents the number of persons of each ability.

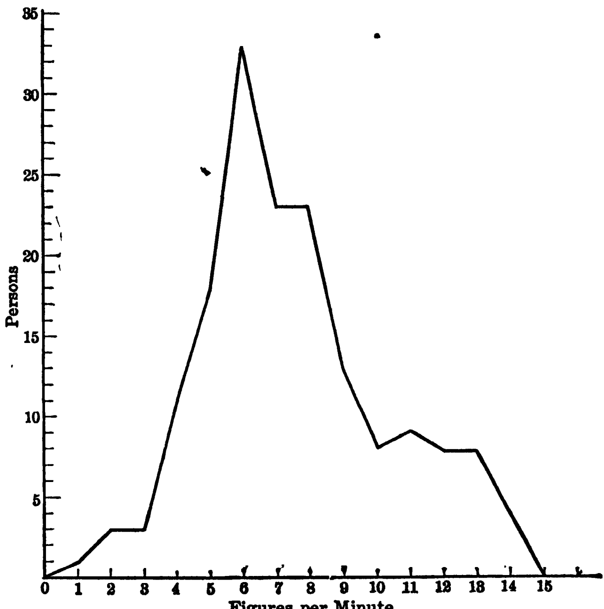

FIG. 9.—Distribution of ability in canceling a certain geometrical figure in a page of figures. Time allowed, one minute. Based on 164 persons.

traits combine in the same person. The differences are qualitative only in the sense that combinations of varying amounts of diverse traits occur.

The most convenient manner in which to represent and determine the amount of variation in a given trait is by means of the distribution curve, or the surface of frequency. The distribution curve is a curve designed to represent how frequently each amount

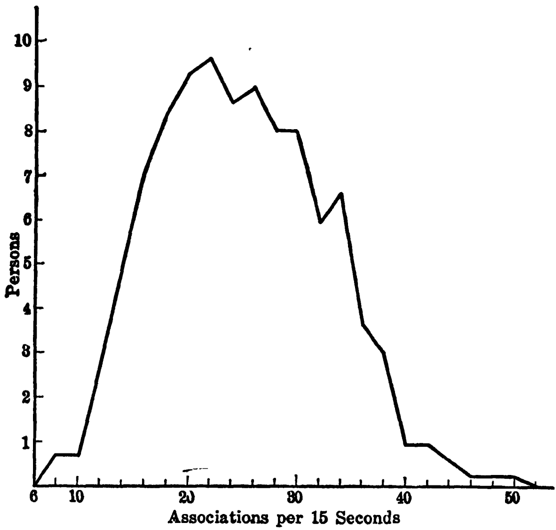

FIG. 10.—Distribution of ability in giving associations in response to a stimulus word. The horizontal axis gives the number of words given in 15 seconds. The vertical axis gives the number of persons of each ability. Based on 135 persons.

or strength of a given trait occurs in a given group of persons. The range of ability from a small amount to a large amount is represented along the base line, or x axis, from left to right, and the number of times each particular ability occurs is represented vertically along the ordinates, or y axis. (See Figures 7, 8, 9, and 10.)

How Wide are the Differences? The investigation of this problem in recent years has brought out the fact that the differences among human beings are very much greater than has commonly been thought. If we measure a group of pupils in a given class or grade, we find that on the average the best pupil is able to do from two to twenty-five times as much as the poorest pupil, or is able to do the same task from two to twenty-five times as well as the poorest pupil. The accompanying table shows the range of differences between the highest and the lowest in a series of tests made upon fifty university students.

TABLE 1

Range of differences between the best and the poorest in a series of mental tests. Based upon the writer's Experiments in Educational Psychology, page 8, which may be consulted for the nature of the tests.

	Best Record	Poorest Record	Ratio
Memory span	8 words	4 words	1 : 2
Memorizing	1 min.	4 min.	1 : 4
E Test	25 sec.	1 min. 30 sec.	1 : 3.6
Er Test	1 min. 30 sec.	3 min. 25 sec.	1 : 2.3
Opposites	30 sec.	1 min. 50 sec.	1 : 1.37
Genus-species	45 sec.	2 min. 5 sec.	1 : 2.8
Addition	31 sec.	2 min.	1 : 3.9
Subtraction	20 sec.	1 min. 30 sec.	1 : 4.5
Average			1 : 3.35

What is the Nature of the Variation? From the general appearance and form of the distribution curves of mental traits, we note that abilities range without break from the lowest to the highest. Our common terminology of dividing groups of persons into various classes as dull, mediocre, and bright, on the assumption that they may be divided into distinct classes with gaps between them, is psychologically incorrect. The fact rather is that all grades of ability, varying by infinitesimally small amounts from the lowest to the highest, are found in the human species.

The next conspicuous feature about the nature of the distribution of mental abilities is the general shape of the curve. This indicates that the large majority of individuals cluster about the center. In the accompanying illustration it will be noticed that if the entire range of abilities is divided along the base line into three equal sections so that we may designate the one at the right as the superior section, the one in the middle as the medium section, and the one at the left as the inferior section, we find that approximately two-thirds, or 66% of all persons fall into the middle third; one-sixth or 17% fall into the superior one-third, and the remaining one-sixth, or 17% fall into the inferior one-third of the range of abilities. In other words, the normal distribution curve is a symmetrical, bell-shaped figure, having its mode in the center and dropping at first rather gradually, then very rapidly and finally very slowly. The statement attributed to Lincoln that "God must have loved the common people because He made so many of them" is psychologically true. If the middle third of the entire range of abilities represents the common people, then two-thirds of all persons are common people.

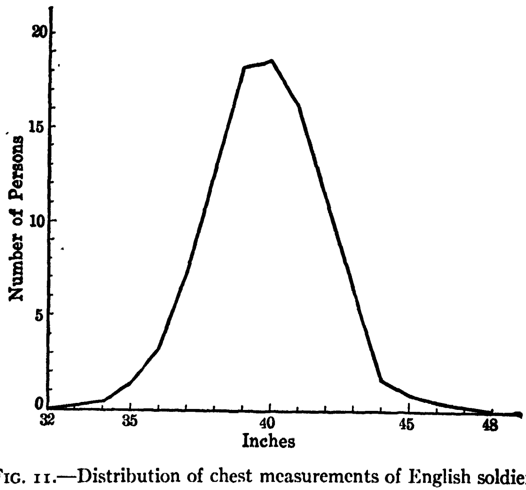

FIG. 11.—Distribution of chest measurements of English soldiers.

The third interesting fact to be noted is that psychological and biological traits vary universally in the same manner in conformity with the normal, bell-shaped curve. Note for example the distribution of such biological traits as chest measurement, height, girth of head, and so forth, as represented in the accompanying graphs, Figures 11, 12, and 13. The number of men who are extremely tall or extremely short is very small, and the number less tall or less short is larger and larger as the median is being approached. This uniformity throughout organic nature is an interesting and significant fact. Apparently nowhere are there traits which are discontinuous so that gaps would exist within the ranges of the traits, nor do we find that, on the whole, traits are distributed

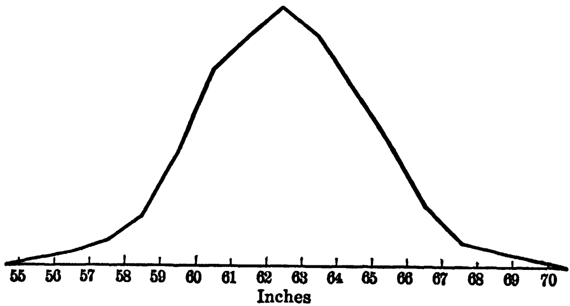

FIG. 12.—Distribution of the height of 1,052 women.

in a skewed manner, so that the great majority of individuals would lie either in the upper or lower range of abilities.

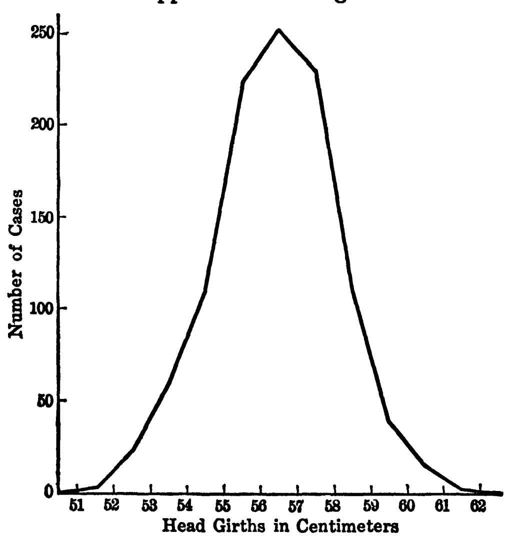

FIG. 13.—Distribution of the head girth of 1,071 boys, 16–19 years of age.

Finally, the variation in both psychological and biological traits occurs apparently according to the law of chance, that is, according to the frequency of occurrence of a chance event. Consequently, on this assumption, the statistical treatment of the distribution of mental abilities becomes subject to the mathematical properties of the probability integral. What we mean by the statement that the variation occurs according to the law of chance may be illustrated in the following manner: If we toss up ten pennies at one time, count the number of heads up and keep a record of it, then repeat the tossing a thousand times and keep a record each time, it will be found that the number of times no heads are up will occur very

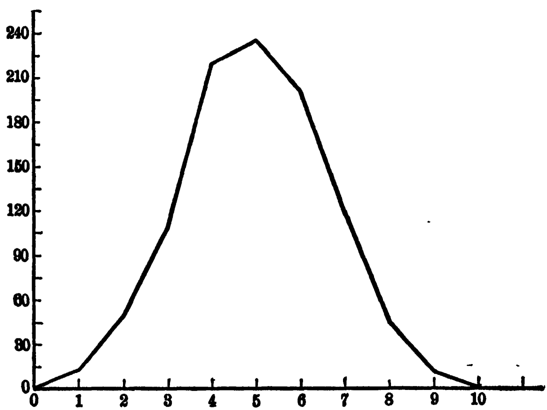

FIG. 14.—Distribution of the number of heads up in tossing ten pennies 1,000 times. The horizontal axis gives the number of possible heads up in each tossing; the vertical axis gives the number of times each number of heads was up.

rarely, likewise, the number of times all ten heads are up will occur very rarely, the number of times one head is up or nine heads are up will occur less rarely, and as you approach from either side toward four, five, and six, the occurrences will be more and more frequent. The actual records of a thousand such tossings are represented in Figure 14. It would seem as though nature, in the production of her creatures, aimed at a target. The largest number of trials strikes somewhere near the bull's-eye, a smaller number strikes within the next circle, and a still smaller number within the next circle, and so on. The correspondence between the actual distribution of abilities and the values of the probability integral is exceedingly useful in permitting statistical treatment of series of measurements of any trait. Figure 15 gives the mathematical or theoretical probability curve.

Variation in Abilities in School Subjects. The differences in

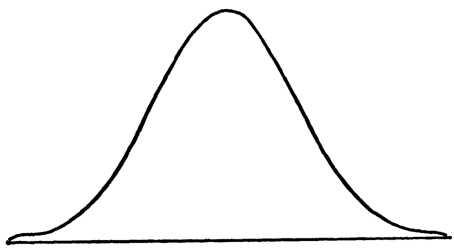

FIG. 15.—The theoretical probability curve.

abilities in school subjects are fully as wide as in special psychological capacities. They are probably due primarily to native

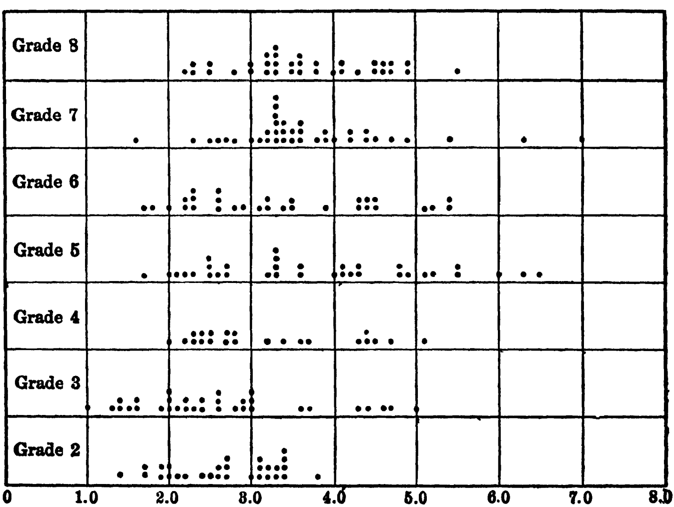

FIG. 16.—Distribution of pupils (in one school) in grades 2 to 8 in reading ability as measured by the author's tests. The horizontal axis represents speed and comprehension combined in terms of speed, i. e., words read per second.

ability rather than to differences in opportunity, training, or environment. Table 2 shows the range of difference in ability

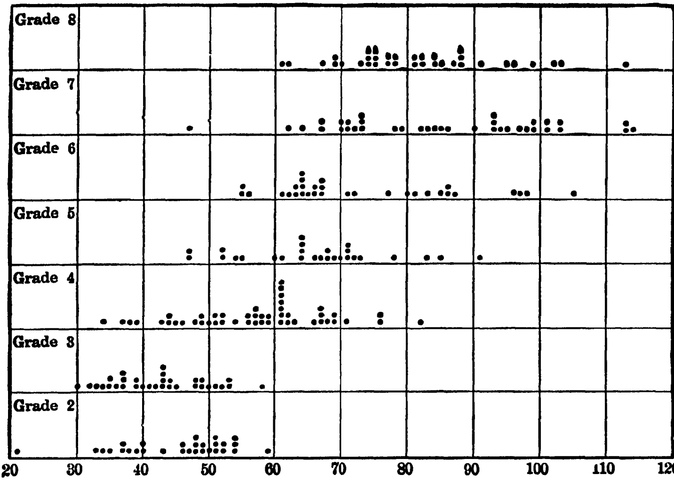

FIG. 17.—Distribution and overlapping of pupils in writing. Speed and quality combined into one score as explained in the author's Educational Measurements. The numbers along the horizontal axis represent speed and quality in terms of speed, i. e., letters per minute. Quality was measured by the Thorndike scale.

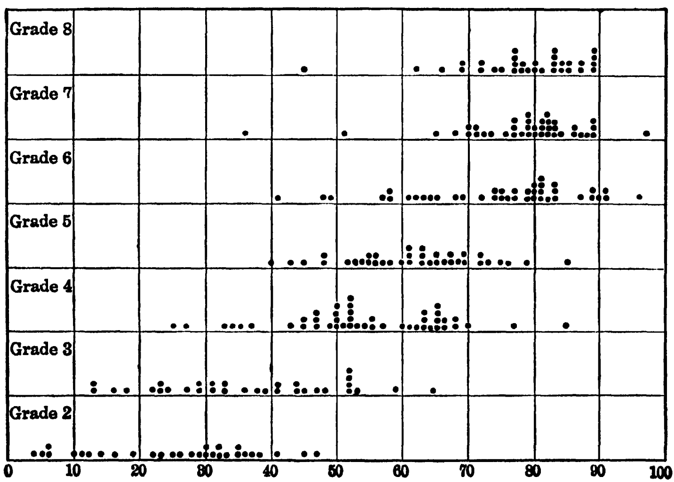

FIG. 18.—Distribution and overlapping of pupils in spelling as measured by the author's test. The numbers along the horizontal axis are the numbers of words spelled out of a list of 100.

in various school subjects as found in a class of 36 eighth-grade pupils. Abilities in reading, arithmetical reasoning, spelling, grammar, and history were measured by the author's tests ('16).

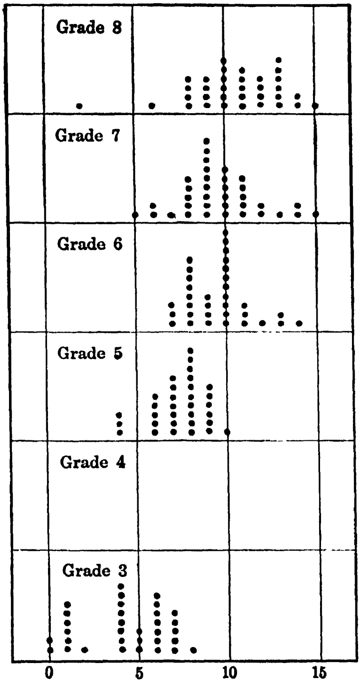

FIG. 19.—Distribution and overlapping of pupils in ability to solve arithmetical problems as measured by the author's Scale A. The numbers along the horizontal axis are the steps on the scale.

Quality of writing was measured by the Ayres scale, the four fundamental operations in arithmetic by the Courtis tests (series B), and composition by the Hillegas scale ('12.)

TABLE 2

Ranges of difference between the best and the poorest in a class of 36 eighth-grade pupils.

	Best	Poorest	Ratio
Reading—Speed	6.6	1.8	1:3.7
Reading—Comprehension	76.	22.	1:3.5
Writing—Speed	108.	57.	1:1.6
Writing—Quality	90.	60.	1:1.5
Arithmetic—Reasoning	15.	2.	1:7.5
Arithmetic—Addition (Rights)	15.	1.	1:15.
Arithmetic—Subtraction (Rights)	17.	2.	1:8.5
Arithmetic—Multiplication (Rights)	17.	1.	1:17.
Arithmetic—Division (Rights)	16.	2.	1:8.
Spelling	90.	45.	1:2.
Composition	70.	30.	1:2.3
Grammar (Scale A)	13.	6.	1:2.2
History	104.	4.	1:26.
Average			1:7.6

In the accompanying diagrams, Figures 16–26, the complete distribution of the abilities of the pupils in each grade in the subjects of reading, writing, spelling, etc., are shown as determined by methods of measurement described elsewhere. These graphs show that the range from the lowest to the highest ability in any given subject within any given grade, is approximately as great as that found for special mental functions referred to in a preceding section. The best pupil in reading or spelling or any school subject is from one and a half to twenty-five times as capable as the poorest pupil. As a result of this wide range of abilities, there appears an enormous amount of overlapping of the abilities possessed by the pupils in other grades in the same school. Thus it will be noted that the best pupil in arithmetical reasoning in the third grade is as capable as the poorest pupil in the eighth grade. All pupils had been tested by the same set of problems. The same statement applies with practically identical details to any school subject. Putting the situation in a little different statement, it has been shown that 60% of the best pupils in any given grade could be exchanged with the 60% of the poorest pupils in the next higher grade, with the result that there would be no change in average ability of the two grades.

The question next arising is this: Granting that the range of ability in any one subject is as large as the results of the tests show it to be, may, however, a given pupil not be two or three years ahead of his grade in arithmetic, two or three years behind his grade in spelling, up to the average of his class in reading, etc., and may he not be placed correctly, after all? The facts, however,

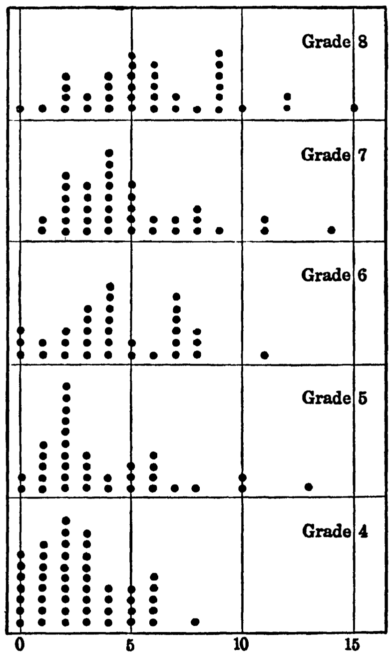

FIG. 20.—Distribution and overlapping in addition as measured by the Courtis test. The numbers along the horizontal axis represent the number of examples done correctly.

seem to be as represented in the accompanying illustration, Figure 27, in which a combined score was obtained for each pupil as follows: In reading and writing in grade 1; in reading, writing, and spelling in grade 2;. in reading, writing, spelling, and arithmetic in grades 3 and 4; and in reading, writing, spelling, arithmetic, language, and composition in grades 5 to 8. Even when the variations in abilities in different subjects possessed by the same pupil are counterbalanced and averaged, the range of abilities and the overlapping is practically as large. It will be noticed, for example, that the best pupils in the second and third grades in these three subjects combined, are almost up to the ability of the poorest pupils in the eighth grade. The fact is that in every eighth grade

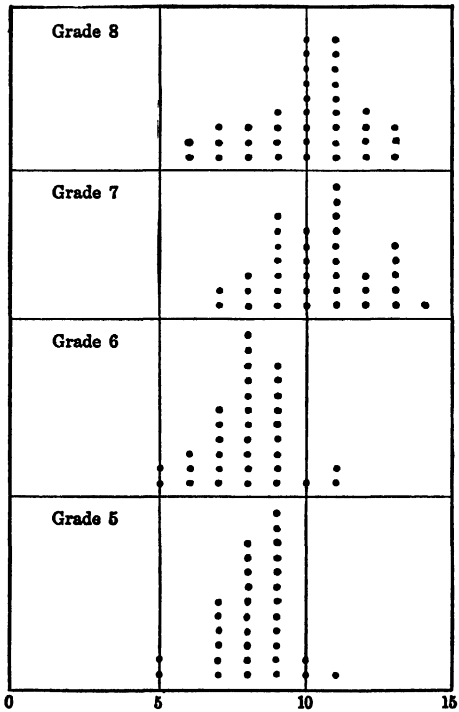

FIG. 21.—Distribution and overlapping in use of correct English as measured by the author's Grammatical Scale A. The numbers along the baseline are the steps on that scale.

one pupil in nine is fully equal in ability to the average ability of the pupils in the second year of high school and could do the work equally well if he had been allowed to go on rapidly enough to be in the second year of high school. Two pupils in every nine are equal in ability to the average pupil in the first year of high school, three of the nine pupils are correctly placed in the eighth grade, two are equal only to the average seventh grader, and one is equal only to the average sixth grader. Thus by proper promotion or classification, one pupil in every nine could save two years in eight, and two pupils in every nine could save one year in eight.

Expressing the same facts in a different form for the school population as a whole, we may say that:

1	pupil	in 100	could finish the 8 grades in	4	yrs. or at	10	yrs. of age.
2	pupils	" "	" " " " "	5	" " "	11	" " "
9	"	" "	" " " " "	6	" " "	12	" " "
21	"	" "	" " " " "	7	" " "	13	" " "
33	"	" "	" " " " "	8	" " "	14	" " "
21	"	" "	" " " " "	9	" " "	15	" " "
9	"	" "	" " " " "	10	" " "	16	" " "
2	"	" "	" " " " "	11	" " "	17	" " "
1	pupil	" "	" " " " "	12	" " "	18	" " "

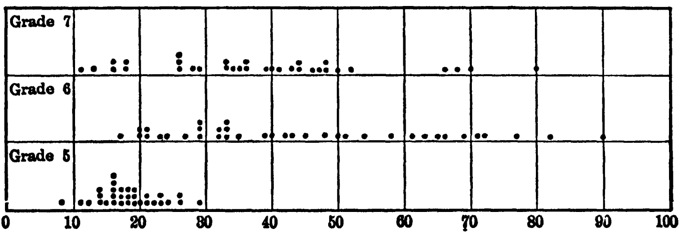

FIG. 22.—Distribution and overlapping in geography as measured by the author's geography test. The numbers along the horizontal axis are the scores in the test. The situation in the case of history is very similar.

The last two groups are composed of pupils so retarded that they probably never would or could complete the elementary school. The variation in ability is so great that the children of any given age are spread out over about nine years of maturity. For example, children ten years old range in ability all the way from fourteen-year-olds to six-year-olds or less, and the numbers of pupils at each age of mentality are approximately those given above. These facts are further borne out by recent tests of intelligence. (See Chapter VII.)

This enormous range of ability and the resulting overlapping of successive grades, is probably the most important single fact discovered with reference to education in the last decade. The import of it is so significant of the situation as it exists in our schools to-day and of the possibilities in the direction of the proper reclassification or

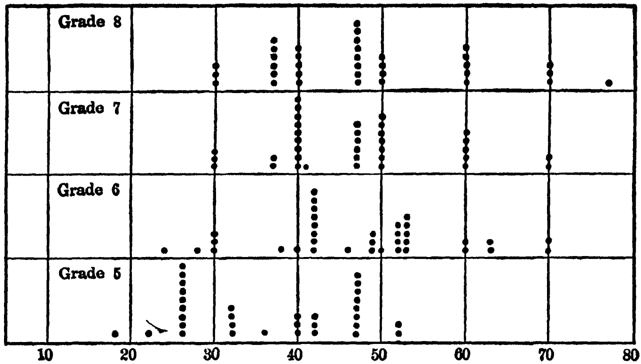

FIG. 23.—Distribution and overlapping in ability to write a composition as rated by the Hillegas scale. The numbers along the base are values on that scale.

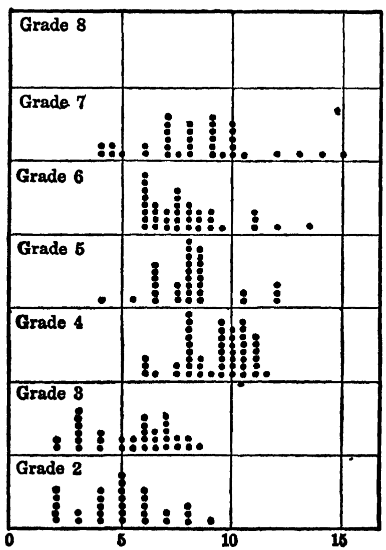

FIG. 24.—Distribution and overlapping in drawing ability. The numbers along the horizontal axis are the units of Thorndike's drawing scale.

readjustment of pupils according to ability that we have scarcely begun to realize how great the differences are or in what manner the readjustments may be made.

Provisions Made in the School for the Variations in Abilities. Experimental work has drawn renewed attention to the possibilities of taking account of the enormous ranges of abilities such as are

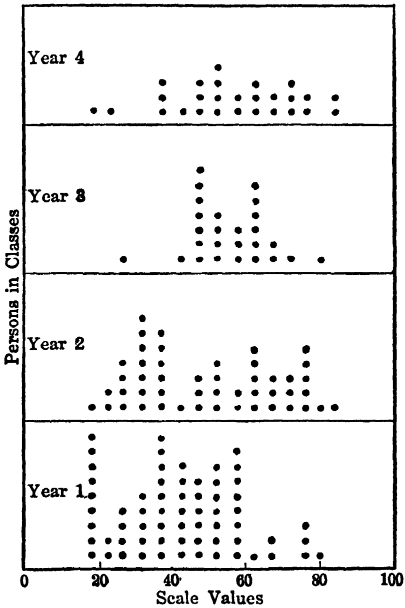

FIG. 25.—Distribution and overlapping of pupils in a high school in ability to write an English composition. The numbers along the horizontal axis are values on the Hillegas scale.

found even in an ordinary class of supposedly homogeneous pupils. To keep an ordinary class of pupils together is no doubt very wasteful in time both for the gifted as well as for the stupid pupils. The gifted must listen to questions and explanations designed chiefly for the benefit of the dull pupils, but which the bright pupils already understand. The dull pupils, on the other hand, waste time by being dragged along too rapidly in the endeavor to keep the bright pupils occupied.

The plans which have been proposed for meeting the varying abilities of pupils fall into two general classes: First, those which attempt to keep the pupils of a given class together but vary the manner of instruction for the pupils of different capacities; second, those which keep the manner of instruction uniform but promote or retard pupils according to their achievements.

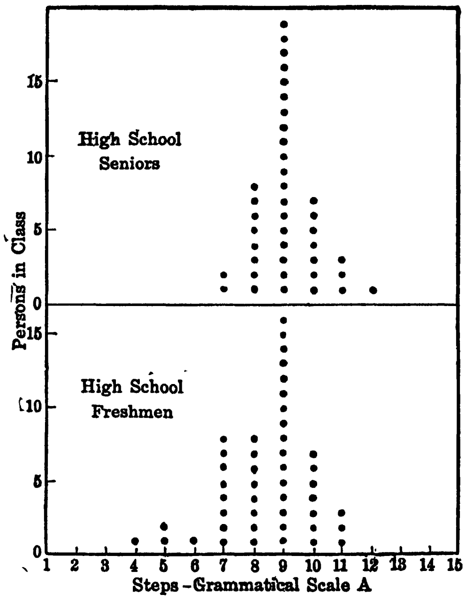

FIG. 26.—Distribution and overlapping of pupils in a high school in ability in discriminating between correct and incorrect English. The numbers along the horizontal axis are the steps on the author's Grammatical Scale A.

The principal schemes of the first general method which have been tried in various schools are known as the individual instruction or Pueblo plan, the monitorial group plan, the extra-work plan, and the supervised study or Batavia plan. The individual instruction plan was employed by Superintendent P. W. Search in Pueblo, Colorado, and consisted in the abolition of all class instruction and the substitution of individual teaching according to the needs of the pupils. The monitorial group plan is carried out by dividing a class into several groups, usually three, according to the abilities of the pupils, and by appointing a monitor for each group from among the members of the class. The extra-work plan consists in having recitation and class instruction chiefly for those who need it, and in assigning additional work to the capable pupils to be done at their desks. The supervised study plan de-

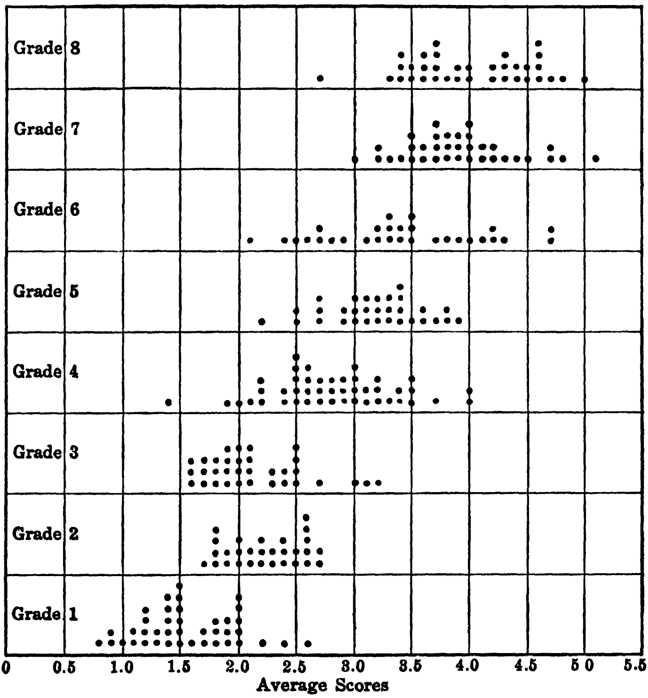

FIG. 27.—Distribution and overlapping of pupils when their attainments in different subjects are averaged. Reported in a thesis by Helen Craig in the library of the University of Wisconsin, 1918.

votes a part of the class period to the usual recitation and instructional work, and the remainder to study done under the supervision of the teacher. Sometimes the class period is considerably lengthened and no home study is done; at other times, the class period is kept at the usual length with some assignments for home study.

These plans have been in operation in various schools during the past thirty years with varying amounts of success or failure. Most of them have been successful when carried out under the immediate supervision of the persons who devised them. The difficulty, however, has usually been that when others have attempted to use them they have not been so satisfactory. Some of the schemes have been objectionable on other grounds also. For example, the individual instruction plan is in part unsatisfactory because it removes a large share of the social stimulus and interaction that is derived from class instruction.

The one type of plan which is being adopted on an extensive scale and is proving to be generally applicable, is some form of the supervised study plan. The methods with which this plan is carried out differ considerably and great care must be taken to avoid formality in the division of the time between recitation and study during the class period and in the order and manner in which the supervision is carried out. A more detailed discussion will be given in the chapter on "How to Study" where this subject properly belongs.

The different schemes coming under the second general proposition, namely, that of keeping the manner of instruction constant and varying the rate of promotion, have been applied widely, and many different plans designed to produce greater flexibility in the rate of promotion have been worked out in various school systems. As illustrations, two plans will be mentioned because they have been in successful operation for many years. In Cambridge, Massachusetts, there has been in operation a plan for some twenty years, in which the work of grades three to eight is laid out in three different courses of study. Pursuit of course A permits the completion of the remaining six grades in six years; pursuit of course B permits the completion of the work in five years; and the pursuit of course C makes possible the completion of the six grades in four years. Transfer from one course to another may take place at any time.

In the St. Louis schools a method of promotion has been in force for a great many years which consists in dividing the school year into four quarters of ten weeks each. Promotion can be made at the end of each quarter. Pupils who have made a grade of excellent may be promoted to the next higher class at the end of any ten-week period, and pupils who have made very low grades or practically failed, must repeat their work beginning with the various ten-week periods.

The effect of this plan in shortening the time of a considerable proportion of pupils is shown in a study made by W. J. Stevens.[1] This investigation shows the length of time required by each of 1,439 pupils in four elementary schools in St. Louis to complete the eight grades.

TABLE 3

The average attendance per grade of 1,439 pupils, graduates, required to complete each of the eight grades. Forty weeks is assigned in the course of study for each grade.

Number of Pupils	Average Number of Weeks to Complete Each Grade	Number of Pupils	Average Number of Weeks to Complete Each Grade
1	17	33	41
2	18	49	42
1	20	29	43
1	21	27	44
8	22	19	45
8	23	20	46
13	24	15	47
17	25	9	48
19	26	5	49
25	27	4	50
46	28	4	51
43	29	2	52
52	30	2	53
83	31	2	54
103	32	1	55
99	33	2	56
109	34	2	57
92	35	2	58
110	36	1	59
87	37	1	60
104	38	2	62
95	39	1	63
87	40	2	70

Median..35 weeks
Total average time of attendance.................288 weeks
To do..320 weeks' work
Double promotions..............................17%
Normal promotions..............................67%
Repeaters......................................16%

It will be noticed from this table that thirteen pupils completed the eight grades in an average of twenty weeks to do forty weeks

[1] Reported in a thesis (1914) in the library of the University of Wisconsin. The study was carried out under the direction of Professor E. C. Elliott.

of work, that is, in half of the prescribed amount of time. In other words, about 1% of the pupils required approximately four years, 6.3% five years, 22.8% six years, 34.6% seven years, 24.9% eight years, 7.6% nine years, 1.7% ten years, and 1.3% eleven to thirteen years to complete the eight grades. These results agree quite closely with the figures suggested on page 39.

Promotion by subjects is a plan adopted in various schools. The program must be arranged so that all grades recite in the same subject at the same period in order that a pupil may do his work with the particular class to which he belongs. For example, a fifth-grade pupil might recite in spelling with a seventh-grade class, in reading with a sixth-grade class, in arithmetic with a fourth-grade class, and so on.

In high school work there is equal need for flexibility in progress. Plans should be devised whereby a class could be divided into three sections, a rapid, a normal, and a slow section. For example, an algebra class, after some early tests, could be divided into three divisions. Section A could easily do the year's work in two-thirds of the year and then pass on to geometry or more advanced algebra or even some other subject. Section B could do the normal work in the year, and Section C could take a year and a third to do the normal year's work, or could cover only two-thirds of the ground in the year and receive only two-thirds credit. Differences in ability are sufficiently great to make possible as much difference in progress as is here indicated. The more capable pupils could easily shorten their high school course by half a year or a year.

A plan of flexible promotion that can be administered successfully has in many respects distinct advantages over any plan which merely varies the instruction for the reason that it allows the capable pupil really to gain the advantage of his ability; because he is able to shorten his elementary school period, which is one of the aims striven for at the present time. The elementary school course is considered too long. Any plan which varies the method of instruction so as to require more work of the capable pupil no doubt occupies the time of these pupils and gives them the benefit of the additional work achieved, but it does not give the pupil the full benefit that he deserves in accordance with his capacities. In practical life the capable man performs several times as much work or makes several times as rapid progress in the same period of time as the incapable man, both having equal opportunities.

Why should not the schools permit progress according to ability and achievement? Greater flexibility in promotion or retardation is an advantage both to the more gifted and to the less gifted pupils. The former will be able to step forward whenever they are ready and the latter will not need to step back so far whenever a part of the work must be gone over again. Promotion once a year works to the disadvantage of both types of pupils. The bright pupils cannot well jump an entire year and so will not be able to progress as rapidly as their abilities warrant, while the slower pupils will have to repeat an entire year when a quarter or half of a year would be sufficient. School progress is determined too much by the calendar and not enough by capacity. The most capable one-third of pupils are advanced too slowly, and the least capable one-third are advanced too rapidly. A saving of half a year or a year on the part of a fourth or a third of the pupils would be of inestimable value to the pupils themselves and to the community at large, either in getting an earlier start in their life work, or, preferably, in securing more advanced and thorough training.

Finally, one of the most important, if not the most important aspect of the principle of progress according to performance, is its appeal to the individual to do the best he can. Few incentives are as strong as the personal impulse of going forward as rapidly as possible and of putting forth the best that is in one. If a child knows that, if he can spell as well as the pupils in the grade above him, he will be put with them, he will be stimulated as he would be in no other way to reach that degree of attainment. Likewise, if he knows that he is likely to be put back to recite in spelling with the pupils of a lower grade if he falls behind, he will put forth his best efforts to hold his own. Dawdling could hardly be encouraged more than it is in many of our schools. Rewards in adult life are more nearly according to ability and performance. The same conditions would work to the advantage of school life.

The schools have given special attention to the backward pupils by organizing separate classes for them and by giving them extra help, but they have given little or no attention to the advanced pupils. Society would be compensated far more for paying at least equal attention to the gifted pupils since they primarily will determine the future progress of mankind. The leaders of society will come from the right end rather than from the left end of the distribution curve. Wisdom would dictate that we devote at least as much care and thought to them, that we surround them with an atmosphere of high aspiration and achievement and stimulate to the full their powers of originality and discovery. This would make for maximum progress based upon ability and performance, not upon birth or social caste.

CHAPTER IV

CORRELATION AMONG HUMAN CAPACITIES

Problem. Any given single trait varies over an enormously wide range among the members of the human race as a whole. The problem, however, before us now is: To what extent is a given amount of any capacity accompanied in general in the same person by equal, larger or smaller amounts of any other ability? To what extent is a good memory in the same person accompanied by an equally good capacity for reasoning or attention or perception or judgment? To what extent is poor or mediocre ability in memory accompanied by poor or mediocre ability in other directions? If all mental abilities were measured on a scale of 0 to 10, the concrete problem would be: To what extent would a memory ability of 7 be accompanied in the same person by a perception ability of 7, or a judgment ability of 7? If it is not accompanied by the same amount of other abilities, by how large or small an amount of any other ability is it accompanied?

Educationally the problem is important and takes the following form: To what extent may we expect pupils, who are excellent, mediocre, or poor in one subject to be excellent, mediocre or poor in other subjects? To what extent is a statement such as the following true in general: "I simply cannot learn languages or mathematics, although I get along very well in my other studies"? To what extent is freedom of electives in studies justifiable on the basis of variation in the combination of capacities in the same individual? To what extent are mental and physical traits correlated? To what extent are abilities similar at different times of life in the same individual? To what extent is ability in childhood or youth a forerunner of ability in adulthood?

Methods of Measuring Combinations of Traits. The extent to which various amounts of abilities accompany one another is measured or expressed definitely by the coefficient of correlation. The value of the coefficient of correlation ranges from 1.00 through 0 to -1.00. A coefficient of correlation of 1.00 means a complete agreement. If the coefficient of correlation between ability in

Latin and ability in German were 1.00, it would mean that the best pupil in Latin would be also the best pupil in German, the second best pupil in Latin would be the second best pupil in German, etc., down to the poorest pupil in Latin who would also be the poorest in German. As the correlation drops farther and farther below 1.00 toward 0, the closeness of this agreement becomes correspondingly less until 0 is reached. If the coefficient of correlation between ability in Latin and ability in German were −1.00, it would mean that the best pupil in Latin would be the poorest pupil in German, the second best pupil in Latin would be the second poorest in German, etc. As the correlation rises above −1.00 toward 0 the reversal becomes less and less until 0 is reached. A coefficient of 0 means that no relationship exists. A pupil might have any amount of ability in one subject and any other amount of ability in the other subject.[1]

The Correlation Among Specific Mental Abilities. The early investigations in this field found surprisingly small correlations even among apparently very similar or closely related capacities. Thus it was thought that a person might have a good memory for words but not for numbers or faces; he might have a keen perception of words but not of letters or geometrical figures and the like. As typical of the earlier results on correlations we may cite a few from Wissler ('01) as follows:

Auditory memory of figures and	visual memory of figures	.29 to .39
" " " " "	auditory " " passage	.04
" " " " "	memory of length of line	.00
" " " passage "	" " " " "	−.07
" " " " "	quickness in naming colors	.03
" " " " "	reaction time	.12
" " " figures "	" "	.17

The significance of these coefficients may be interpreted approximately as follows: A coefficient of 0 means that no correlation exists, and roughly speaking, a coefficient of .30 or less is small and practically means very little agreement. Correspondingly, a coefficient lying between .30 and .50 means a moderate amount of agreement, a coefficient between .50 and .75 means a considerable correlation, while a coefficient above .75 indicates a very close

[1] For methods of computing the coefficient of correlation, consult the author's *Experiments in Educational Psychology*, (1917 Edition) Chapter IV; Whipple, *Manual of Mental and Physical Tests*, Chapter III; Thorndike, *Mental and Social Measurements*, Chapter XI; and Rugg, H. O., *Statistical Methods Applied to Education*, Chapter IX.

relationship and, as it approaches 1.00, indicates practically perfect agreement. From the table, it appears that many coefficients are very low and imply little or no agreement. The coefficients so low as to indicate practically very little correlation are stated to exist between auditory memory of figures and visual memory of figures. This would mean that a person might have a very good memory for figures seen but a poor memory for figures heard.

The difficulty with these coefficients is that they are based upon unreliable and incomplete measurements of the traits concerned. Many of the measurements of the early investigations of correlations were derived from group tests which had been made but once. Measurements thus obtained have been shown by subsequent investigations to be rather uncertain indications of the real amount of a given trait possessed by the individual. In order to obtain a fairly accurate measurement of a given capacity, it is necessary to repeat several times under favorable circumstances the measurement of the trait in question. You cannot obtain anything like an accurate measure of any amount or quantity by a single measurement made under distracting conditions. If we should wish to measure the memory capacity of a given individual, we should not consider the result very trustworthy if a single test were given consisting of eight lines of poetry learned in three or four minutes. We ought at least to repeat the test with several similar passages, preferably on different days, and derive therefrom an average. This is, in fact, the sort of procedure that has been followed in subsequent researches. Inaccurate measurements, as indicated by recent analyses of correlations, tend to reduce very materially the computed coefficients below the actual amounts of correlation.

Recent researches have shown that among many traits quite close, and among other traits very close, correlations exist. An investigation by J. A. Stevenson ('18) showed remarkably close correlations between various types of sensory discrimination. The plan of the investigation consisted in making extensive and repeated measurements of discrimination in lengths of lines, in intensities of sound, in degrees of brightness, in shades of gray and in pressures on the end of the first finger. The correlations computed on the basis of these measurements with ten persons were as follows:

TABLE 4. After Stevenson ('18).

Lines and intensity of sound	.90
Brightness and intensity of sound	.90
Pressure and intensity of sound	.36
Pressure and lines	.39
Lines and brightness	.92
Pressure and brightness	.41

A similar investigation in the field of memory, conducted by Miss N. F. Bennett ('16), showed on the basis of numerous and repeated tests with nine subjects fairly close correlations among the capacities to remember various kinds of material such as syllables, numbers, nouns, prose, and faces, between visual and auditory presentations of the material, and between mediate and immediate learning. Her conclusions are stated thus:

"1. There is a high correlation between mediate and immediate retention if a sufficient number and variety of measurements for each type of memory are taken, and the results amalgamated to determine ranks.

"2. There is a high correlation between the memory span, or immediate retention for disconnected materials, and the ability to learn the same."

Hollingworth made a study to determine the increase in the coefficients of correlation among six different capacities with the increase in the number of measurements made upon each capacity. His results are set forth in the following table. They indicate a very marked rise in the coefficients with the increase in the number of tests.

TABLE 5

The average correlation of each test with all others at various points in the curve of practice. After Hollingworth ('12).

Trial	Adding	Opposites	Color Naming	Discrimination	Co-ordination	Tapping	Final Average
1	.19	.10	.15	−.07	−.15	.17	.065
5	.41	.26	.15	.35	.21	.32	.280
25	.50	.35	.43	.27	.03	.35	.320
80	.55	.43	.53	.31	.18	.34	.390
205	.48	.62	.61	.35	.34	.52	.490

As an illustration of a series of correlations among special mental functions based upon measurements repeated several times but not as frequently as those in the preceding tables, we may cite the coefficients obtained by Simpson. These coefficients are unusually high because they are based upon tests performed on two extreme groups of subjects, the one a highly intelligent group and the other distinctly unintelligent.

TABLE 6

Correlations among certain mental abilities. After Simpson ('12).

	Ebbinghaus Test	Hard Opposites	Memory of Words	Easy Opposites	A-Test	Memory of Passages	Adding	Geometrical Forms	Learning Pairs	Completing Words	Drawing Lines
	1	2	3	4	5	6	7	8	9	10	11
1. Ebbinghaus test											
2. Hard opposites	92										
3. Memory of words	92	92									
4. Easy opposites	75	81	68								
5. A-Test	68	76	70	71							
6. Memory of passages	91	86	89	69	60						
7. Adding	71	74	56	70	67	66					
8. Geometrical forms	54	64	67	54	94	60	44				
9. Learning pairs	72	72	82	43	44	63	46	40			
10. Completing words	50	70	51	50	84	38	77	61	34		
11. Drawing lines	26	25	06	53	27	12	27	30	04	17	
12. Estimating lengths	52	55	59	56	57	58	17	35	54	22	55

Burt obtained the following correlations (Table 7) from a variety of tests of specialized mental functions made upon forty-three pupils. The test designated as dotting was regarded as a measure of voluntary attention; the tests designated as spot pattern, mirror and memory were designed to measure memorial and associative capacities; the tests called alphabet and sorting referred to sensorimotor capacities; dealing and tapping to motor functions; and the remainder to sensory discrimination.

TABLE 7

After Burt ('09).

	Dotting	Spot Pattern	Mirror	Memory	Alphabet	Sorting	Dealing	Tapping	Sound	Lines	Touch	Weight
Dotting................												
Spot pattern............	.71											
Mirror..................	.92	.75										
Memory..................	.53	.62	.38									
Alphabet................	.84	.89	.63	.73								
Sorting.................	.87	.57	.67	58	.91							
Dealing.................	33	.53	.05	.12	.60	.54						
Tapping.................	.78	.57	.74	.40	.80	.89	.57					
Sound...................	.65	.47	.66	.23	.51	—.18	.17	.48				
Lines...................	.52	.67	.55	.16	.35	.19	.26	.36	.09			
Touch...................	.30	—.04	.38	.06	.09	—.06	—.08	.10	.12	.23		
Weight..................	.16	.18	.30	.14	.10	—.05	.17	.42	.29	.00	49	—

The import of the researches up to the present time seems quite certainly to prove that the higher mental capacities are on the whole rather closely correlated. The coefficients lie for the most part above .50, and some of them reach up to .80 and .90. The same statement holds approximately for sensory capacities among themselves and also probably for motor capacities among themselves. The cross-correlations among traits from these three levels is, so far as we can judge at the present stage of our knowledge, lower than among the traits within a given level. This seems to be particularly true of the correlation of motor capacities with intellectual capacities.

Correlations Among Abilities in School Subjects. The development of knowledge concerning this aspect of our problem has had a history similar to that of the special mental functions. The early correlations among abilities in school subjects were computed upon relatively uncertain data. About 1903, coefficients obtained by various investigators, were as follows:

TABLE 8

Summary of coefficients of correlation between abilities in high school subjects as reported up to about 1903. (Thorndike '03, pp. 26, 30–31).

B = After Burris, based on nearly 1,000 pupils.
P = After Parker, based on 245 pupils.
Br = After Brinckerhoff, Morris, and Thorndike.

		English	History	Science	Algebra	Drawing	German	French	Latin
History	B...........	.40							
	P...........	.62							
	Br..........	.41							
Science	B...........	.41	.40						
	P...........	.58	.56						
	Br..........	.26	.61						
Algebra	B...........								
	P...........	.55	.38	.40					
	Br.........								
Drawing	B...........								
	P..........	15	10	.33	.20				
	Br..........	.20	.16	.30					
German	B...........								
	P...........	.65	.49	.62	.52	.06			
	Br..........	.30	.42	.58					
French	B...........								
	P...........	.49	.58	.48	.68	.30	.33		
	Br..........								
Latin	B...........	.48	.43	.44					
	P...........	.62	.43	.54	.54	.01	.38		
	Br.........	.50	.44	.35		.40			
Mathematics	B...........	.39	.33	.41					.40
	P...........								
	Br..........	.09	.26	.07		.02	.48		.31

In the case of grammar school subjects, A. G. Smith (Thorndike '03 p. 13), computed the following correlations:

English and Mathematics....................	.39
" " Geography......................	.43
" " Drawing........................	.15
Mathematics and Geography..................	.36
" " Drawing....................	.14
Geography " "	.12

These coefficients for the most part indicate only a moderate amount of correlation. Thus Thorndike interpreted them in 1903 by the following statement: "For our purpose the most striking thing about these figures is their small amount. It is safe to say that in a grammar or high school student a deviation from the average ability in any one subject implies by and large a deviation in any other not more than half as great. The most talented scholar in one field will be less than half as talented in any other, The most hopeless scholar in one field will in another be not so very far below mediocrity." ('03, pp. 37–38).

The coefficients here quoted were based usually upon marks of a single teacher in any given subject. Recent studies have called attention to the unreliability of marks and the differences in standards of marking employed by different teachers. See Chapter XXII. This necessarily produces a considerable reduction in coefficients based upon them.

A computation based upon the average mark of each pupil in each subject in grades five to eight yielded the following coefficients (Table 9):

TABLE 9

Correlations among abilities in school subjects. After Starch ('13).

Arithmetic and language	85
" " geography	.83
" " history	.73
" " reading	.67
" " spelling	.55
Language and geography	.85
" " history	.77
" " reading	.83
" " spelling	.71
Geography and history	.81
" " reading	.80
" " spelling	.52
History and reading	67
" " spelling	.37
Reading and spelling	.58

These coefficients are almost twice as high as those previously quoted and represent very close correlations. They would warrant the interpretation that the pupil who is good, mediocre, or poor in a given subject, is good, mediocre, or poor to very nearly the same, but not equal, degree in all other subjects so far as his abilities are concerned. Such lack of agreement as does exist is due probably to a difference of interest and industry on the part of the pupil in different subjects at different times and to a real difference in abilities in the various fields. Thus spelling ability correlates apparently less closely with ability in other subjects than abilities in these other subjects correlate among themselves. The up-shot of the whole problem concerning the variation in the combination of traits, or the extent to which different amounts of mental traits accompany one another, may fairly be stated as follows:

First, no negative correlations exist either among the abilities involved in school subjects or among the special mental functions measured by special tests. Popular and "short-story" psychology is false in the assumption and description of antagonisms of mental traits. They apparently do not exist among desirable and useful traits. Advice, given to prospective wives, such as "if he is good-natured, he may be lazy; if he is scholarly, he may be cold; if he is thrifty, he may be stingy; if he is generous he may be wasteful," may produce caution, but it is not true psychology. Good-natured men are probably on the whole no more lazy than ill-natured men are, and scholarly men are probably on the whole no more cold-hearted than stupid men are. In fact the opposite is more likely to be true. And such statements as "Johnny is very bright in reading, but he simply cannot get arithmetic" is a soothing salve for the feelings of parents, but not apt to be sound psychology.

Most of the opinions of students who state that they "simply cannot get" mathematics or language or history are in part probably due to a relatively small discrepancy in abilities, that is, to somewhat less ability in mathematics, language or whatever the subject may be; but to a larger extent they are illusory, because, when the actual facts are obtained or when more careful measurements of the abilities in various directions are made, the abilities correlate much more closely than the student's statements would lead one to believe. As a concrete example the following case of a college freshman, brought to the author's attention, may be cited: The student claimed that he had always had great difficulty in learning foreign languages but that other subjects were easy for him. He stated that in high school he never was able to obtain a grade in languages higher than about 75 but that in other subjects his grades were always high, as high as 95. Since his trouble seemed to be language it was thought that he might have a defective memory or an abnormal type of imagery. Some memory and imagery tests revealed the fact that he had normal memory and imagery of average ability. This at once led to an inquiry into his actual high school record to ascertain his grades. The various grades for any given subject are final grades in different courses as follows: English, 83, 80, 78, 81; History, 88, 75, 83; mathematics, 80, 87, 77; science, 87, 87; Latin, 77, 79, 75; German, 75, 75. When these marks are compared there is little or nothing to explain. His marks in Latin and German were somewhat lower than in other subjects, which is probably largely explained by his own statement that he "hated" languages, but they were not much lower on the whole. The highest grade in any subject was in the first year of history, 88, but he also had a grade of 75 in the second year of history and 77 in the 3rd course in mathematics. There was no grade of 95 in the entire list. This record was corroborated by the grades which he received at the end of the first eight weeks of his freshman year in college: Spanish, Fair; Geology, Fair; English, Fair; Mathematics, Fair; History, Poor. His abilities are pretty uniformly mediocre in all respects.

Exceptions do occur such as that of a boy seventeen years of age in the second year of the high school who was able to carry his work satisfactorily, but was able to read no more fluently, either orally or silently, than the average pupil can at the end of the first grade. He was a normally intelligent boy. Such cases occur perhaps once among one or two hundred pupils, and may be regarded as abnormal.

Second, intellectual and scholastic abilities are for the most part closely correlated. Barring certain exceptions, which are rarer than is generally supposed, abilities are combined in fairly similar amounts. Intercorrelations between the different levels, intellectual, sensory, and motor, seem to be smaller and in some traits, practically zero. Some of the motor abilities, such as handwriting, have practically no correlation with intelligence or general mental abilities.

The wider bearing of the facts about the combinations of mental capacities, together with the distribution of mental traits according to a continuous, bell-shaped curve discussed in the preceding chapter, are deeply significant for the problem as to whether there are distinct mental types. Mankind apparently cannot be divided into three or four separate types. The ancient classification of temperaments into sanguine, choleric, melancholic, and phlegmatic, may be conveniently analogous to the four seasons of the year, spring, summer, autumn, and winter respectively, but there are no mental types that correspond to such superficial characteristics and none that are marked off sharply or even vaguely from one another. If all members of the human race were to be exhibited in a distribution curve whose base line represented from left to right different amounts of "sanguine-melancholic, or choleric-phlegmatic" natures, the curves would in all probability not be a series of four distinct curves separated from one another, nor even possess four modes with depressions between them, but would very likely be single continuous distribution surfaces of the usual normal form with one mode. The human beings who even remotely approach any one type are very rare. The rule is that each person possesses more or less of all different traits, and within certain limits, roughly similar amounts of the various traits. Persons in whom the divergences are large are the exceptions rather than the rule.

Correlation between Special Mental Capacities and General Intelligence. So far as definite data are available on this point, the inference may be drawn that many special mental functions are correlated anywhere from moderately to very closely with general intelligence. Men of intelligence have, on the whole, keen powers of perception, observation, and attention, remarkable retentiveness, exceptionally rapid and varied association processes, as well as unusually incisive powers of analysis and soundness of judgment. We may note here in passing, by turning to Chapter VII, the amounts of correlation of certain capacities with general estimated intelligence as found by Simpson, Burt, and others.

The usefulness of the facts that many specific mental capacities are reliable symptoms or essential constituents of general intelligence will be particularly important in the future in the development of tests and methods of measuring intelligence. The value of this to mankind, not only in education but in all fields of human endeavor, can hardly be foretold at the present time. Further consideration will be given to it in a later chapter.

Correlations between Mental and Physical Traits. In the case of adults, the correlations between mental abilities and such physical characteristics as height, weight, size of head, lung capacity, or strength of grip, are either very low or zero. In the case of children, the situation is somewhat different. B. T. Baldwin made an elaborate study of 861 boys and 1,063 girls in the University of Chicago elementary and high school, the F. W. Parker school of Chicago, and the Horace Mann School of Columbia University. Measurements of various physical characteristics were obtained at yearly and half-yearly intervals on two groups of pupils. One group was followed continuously through the ages from six to twelve, and the other from twelve to eighteen. A parallel comparison between the physical measurements and the school records of the same pupils was then made. From these results, Baldwin has derived the following conclusion:

> Taller, heavier children mature physically in advance of the shorter, lighter ones. Those whose physiological age is accelerated complete the last grade of the elementary school at 12 years, 9 5/6 months of age with an average of 84.3%, and those below average or of retarded physiological development, complete the elementary school work at 13 years 7 4/13 months of age, with an average of 81.7%. (Bulletin of Bureau of Education No. 581, 1914. Page 82.)

Correlations Between Early and Later Mental Abilities. The problem here is, to what extent will a given pupil maintain his record of excellence, mediocrity, or stupidity all through his educational career or all through his life? Will the pupil who has high, medium, or low ability in the elementary school also have high, medium, or low ability in high school and in college? The first extensive study in this field was made by W. F. Dearborn ('09) who traced through the high school and through the university the scholastic records of various groups of students, varying in size from 92 to 472, and coming from eight large and four small high schools in Wisconsin. He divided the pupils into four quartiles according to their marks in high school, and then ascertained to what extent the pupils remained in the same quartiles during their university course. His records showed that the pupils maintained the same records with remarkable consistency. He states his conclusion in the following words:

> We may say then, on the basis of the results secured in this group (472 pupils) which is sufficiently large to be representative, that if a pupil has stood in the first quarter of a large class through high school, the chances are four out of five that he will not fall below the first half of his class in the university. . . . The chances are but about one in five that the student who has done poorly in high school—who has been in the lowest quarter of his class—will rise above the median or average of the freshman class at the university, and the chances that he will prove a superior student at the university are very slim indeed. . . . The

Pearson coefficient of correlation of the standings in the high schools and in the freshman year, for this group of 472 pupils is .80. . . . A little over 80% of those who were found in the lowest or the highest quarter of the group in high school are found in their respective halves of the group throughout the university. . . . Three-fourths of the students who enter the university from these high schools will maintain throughout the university approximately the same rank which they held in high school.

F. O. Smith made a similar study of 120 students entering the College of Liberal Arts at the University of Iowa. He traced their records from high school through the entire university course and found almost the same situation. Expressed in terms of coefficients of correlations, the results were as follows:

TABLE 10

Correlations. After Smith. ('12.)

H. S. average and Univ. Freshman Average	.48
H. S. Average and Univ. Sophomore Average	.39
H. S. Average and Univ. Junior Average	.47
H. S. Average and Univ. Senior Average	.28
1st and 2nd Year High School	.77
1st and 3rd Year High School	.67
1st and 4th Year High School	.66
University Freshman and Sophomore	.73
University Freshmen and Junior	61
University Freshmen and Senior	.45

T. L. Kelley compared the marks of 59 pupils as they passed from grade five up into the first year of the high school. The extent of the agreement of their records in successive years is shown in the following coefficients of correlation:

TABLE 11. After Kelley. ('14).

Correlation between marks in the grades and marks in the first high school year.

First Year of High School and 7th Grade	.72
First Year of High School and 6th Grade	.73
First Year of High School and 5th Grade	.53
First Year of High School and 4th Grade	.62

He then states:

"The net conclusion which may be drawn from these coefficients of correlation is that it is possible to estimate a person's general ability in the first year (H. S.) class from the marks he has received in the last four years of elementary school with accuracy represented by a coefficient of correlation of .789, and that individual idiosyncrasies may be estimated, in the case of mathematics and English, with an accuracy represented by a coefficient of correlation of .515. . . . Indeed, it seems that an estimate of a pupil's ability to carry high school work when the pupil is in the fourth grade may be nearly as accurate as a judgment given when the pupil is in the seventh grade."

A study of the permanency of interests was made by Thorndike ('12) by comparing the relative strength of interests and abilities during each of three periods of a person's school career, during the elementary school, high school, and college. These comparisons were made by asking one hundred individuals to estimate in retrospect, their relative interests and abilities in mathematics, history, literature, science, music, drawing, and manual work. Such data are necessarily subject to the errors of memory and judgment, but they are practically the only results available so far as strength of interests is concerned. Thorndike inferred from these estimates that early interests are not passing whims, but rather prophetic, with a fair degree of certainty, of later interests and abilities. He concludes that "A correlation of .60 or .70 seems to be approximately the true degree of resemblance between the relative degree of an interest in a child of from ten to fourteen and the same person at twenty-one." "Interests are shown to be not only permanent but also symptomatic to a very great extent, of present and future capacity or ability. Either because one likes what he can do well, or because one gives zeal and effort to what he likes, or because interest and ability are both symptoms of some fundamental feature of the individual's original nature, or because of the combined action of all three of these factors, interest and ability are bound very closely together. The bond is so close that either may be used as a symptom for the other almost as well as for itself. The importance of these facts for the whole field of practice with respect to early diagnosis, vocational guidance, the work of social secretaries, deans, adviser, and others who direct students' choices of schools, studies and careers is obvious."

The impression gained from all these investigations is that human nature is not a medley of capricious capacities which vary from year to year, but rather a fairly consistent combination of abilities throughout life.

CHAPTER V

SEX DIFFERENCES

Educational Significance of Sex Differences. If we may judge fairly at the present time concerning the nature and amounts of differences between the sexes in mental characteristics, it would seem that the differences are so small in native intellectual abilities that they are almost wholly negligible in the education of boys and girls. That boys and girls ought to be educated differently may very probably be desirable, but for reasons other than differences in ability. The professional, business, and domestic life of men and women makes it necessary to have different training for boys and girls. But so far as the native abilities involved in school work are concerned, boys and girls might as well pursue the same courses from the first day of school to the last.

Popular vs. Scientific View of Sex Differences. Probably more fallacious psychology of sex has been spread abroad by novelists and journalists than has been disseminated on any psychological question of popular interest. Occasional and extreme differences in individuals of either sex have been seized upon and exaggerated by descriptive phraseology and represented as though they were the normal divergences between men and women. Up to less than two decades ago, there was practically no scientific knowledge of the nature of sex differences available, and the statements of popular beliefs about such differences were hardly exaggerated by the sort of differences implied in the Sanscrit myth of the creation of woman.

"In the beginning, when Twashtrai came to the creation of woman, he found that he had exhausted his materials in the making of man, and that no solid elements were left. In this dilemna, after profound meditation, he did as follows: He took the rotundity of the moon, and the curves of the creepers, and the clinging of tendrils, and the trembling of grass, and the slenderness of the reed, and the bloom of flowers, and the lightness of leaves, and the timidity of the hare, and the vanity of the peacock, and the clustering of rows of bees, and the joyous gaiety of sunbeams, and the weeping of clouds, and the fickleness of the winds, and the softness of the parrot's bosom, and the hardness of adamant, and the sweetness of honey, and the cruelty of the tiger, and the warm glow of fire, and the coldness of snow, and the chattering of jays, and the cooing of the kokila, and the hypocrisy of the crane, and the fidelity of the chakrawaka, and then compounding all these together, he made woman and gave her to man. But after one week, man came to him and said: Lord, this creature that you have given me makes my life miserable. She chatters incessantly and teases me beyond endurance, never leaving me alone; and she requires incessant attention, and takes all my time up, and cries about nothing, and is always idle; and so I have come to give her back again, as I cannot live with her. So Twashtrai said: Very well; and he took her back. Then after another week, man came again to him and said: Lord, I find that my life is very lonely since I gave you back that creature. I remember how she used to dance and sing to me, and look at me out of the corner of her eye, and play with me, and cling to me; and her laughter was music, and she was beautiful to look at, and soft to touch; so give her back to me again. So Twashtrai said: Very well, and gave her back again. Then after only three days, man came back to him again and said: Lord, I know not how it is; but after all I have come to the conclusion that she is more of a trouble than a pleasure to me; so please take her back again. But Twashtrai said: Out with you, Be off. I will have no more of this. You must manage how you can. Then man said: But I cannot live with her. And Twashtrai replied: Neither could you live without her, and he turned his back on man, and went on with his work. Then man said: What is to be done? For I cannot live either with or without her. (Thomas, *Source Book of Social Origins*, p. 512.)

Such popular beliefs have been in part justified by the probability that many obvious differences are due to the work, and the resulting variation in experience and environment, of women as contrasted with those of men. Thus men know more about business, politics, current events and machines because their occupations bring them much more in contact with these things; but it does not follow that women could not, or would not, know as much about them if their occupations were as much concerned with them. Women know more about cooking, social events, and household utensils because their occupations bring them much more in contact with them; but it does not follow that men could not, or would not, acquire as much knowledge or skill in these directions if their occupations required it.

The differences between the sexes are probably quantitative rather than qualitative. Both men and women have the same reflexes, instincts, and capacities with the exception of certain aspects of the sex instinct. These are probably similar in the main and differ chiefly in their manner of expression. The differences due to sex life and the rearing of children, with the consequent differences in occupations and experiences, will account for many of the superficially observable differences between men and women.

What are the differences that have been scientifically measured and compared? In order to produce a complete picture of mental differences between men and women it would be necessary to measure each trait in a very large number of persons and to compare the measurements with regard to both the averages of the abilities and the manner of the distribution of each ability. This has been done in part only with a few traits and only upon small groups of persons.

Differences in Average Amounts of Mental Abilities. There are two methods by which abilities of two groups may be compared. Either we may state the actual average or median of each group, or we may state how many members of one group reach or exceed the average or median of the other group. The latter method is preferable in many respects to the former in that it makes possible a comparison of groups of various sizes and indicates the relative differences more nearly true to fact. The two methods may be illustrated in the case of a memory test consisting of the oral presentation of ten words at the rate of one word per second and of asking the subjects to record immediately the number of words remembered. He may then state that the number of words remembered on the average by men was 6.9 and by women 7.2. Or we may state that 43.6% of men reached or exceeded the median of the women. The latter method of comparison represents probably more true to life the amount and kind of difference or similarity that actually exist. The differences, hastily inferred from a comparison of averages only, would lead to the conclusion that in regard to memory women are distinctly superior to men. The implication would be that all women have a memory superior to that of men, whereas the fact is that the number of women having a memory superior to that of men is really small and that, in these few women, memory is better only by a very small shade. If 43% of men reach or exceed the median of women, it means that if the 7% of women having a slightly superior memory were omitted, the remaining 93% of the women would have a memory ability identical with that of the men. A difference of 7% in the distributions between two groups is represented by the curves in

Figure 28. The difference is so small that the groups could hardly be distinguished.

By the method of amounts of overlapping in the distribution of one group over the other, the following results have been obtained from students in the University of Wisconsin in a series of tests on memory as just stated, on perception consisting in the cancellation within one minute of as many of a certain geometrical figure as possible, on motor ability consisting in tapping with a pencil upon a card as rapidly as possible for thirty seconds, and on mental addition as described elsewhere.[1]

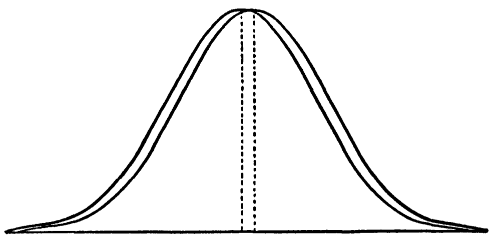

FIG. 28.—Distribution curves representing a difference of 7% between the medians of the two groups.

TABLE 12

Percentage of men reaching or exceeding the median of women.

Perception of geometrical forms......	193 men	200 women	54.5%
Memory of words.................	55 men	77 women	43.6%
Motor ability......................	25 men	50 women	72.0%
Mental addition....................	21 men	46 women	66.7%

In the interpretation of these percentages of overlapping it must be remembered that if 50% of one group reaches or exceeds the median of the other, it means of course that the two groups are identical in ability and distribution. If the percentage of men reaching or exceeding the median of the women is over 50% it means that the men are superior by the number exceeding 50%.

Helen Thompson Woolley made a series of tests as indicated in the following table upon twenty-five men and women at the University of Chicago, on the basis of which Thorndike has computed the following percentages of men reaching or exceeding the median for women:

[1] *Experiments in Educational Psychology*, revised edition, chapter 16.

TABLE 13

Percentages of men reaching or exceeding the median of the women. After Woolley as computed by Thorndike ('14, III, p. 178).

Reaction time	68%	
Tapping	81%	
Sorting cards, speed	14%	
Sorting cards, accuracy	44%	
Thrusting at target	60%	
Drawing lines	72%	
Threshold of pain	46%	
Threshold of taste	34%	(22)
Threshold of smell	43%	
Lifting weights	66%	
Two-point discrimination	18%	(43)
Memory (syllables and learning)	32%	(46)
Ingenuity	63%	

In a similar comparison made on the basis of 100 boys and 100 girls from results obtained by Gilbert, the percentage of boys reaching or exceeding the median of girls was as follows:

TABLE 14

Percentages of boys reaching or exceeding the median of the girls. After Gilbert ('94) as computed by Thorndike ('14, III, p. 182).

	to 14 years	15 to 17 years
Discrimination of weights	48%	58%
" " colors	39%	58%
Reaction time	57%	76%
Resistance to size-weight illusion	55%	68%
Rate of tapping	64%	73%

Thorndike ('14, III, p. 183) reports measurements in which the comparison of the percentages of boys reaching or exceeding the median of girls for persons 8 to 14 years old, were as follows:

TABLE 15

Associative tests, opposites, addition, multiplication, etc.	48%
Perception, A-test, etc.	33%
Memory of words	40%

The writer has made comparisons in the case of school subjects on the basis of abilities measured by means of tests and scales. Speed of writing was measured in terms of letters written per minute. Quality was rated by the Thorndike scale. Attainments in the remaining subjects were measured by the author's tests in these fields. The following percentages of boys, reaching or exceeding the median of the girls, were obtained:

TABLE 16

Speed of handwriting,	about 1100	boys and	1100	girls	47%
Quality of handwriting,	" 1100	" "	1100	"	39%
Arithmetical reasoning,	" 1250	" "	1250	"	60%
History,	" 429	" "	526	"	72%
Geography,	" 447	" "	472	"	48%

Figures of a similar sort computed by Thorndike ('14, III, p. 183) on the basis of teachers' marks showed the following percentages of boys reaching or exceeding the median of girls:

TABLE 17

	High school pupils
English	41%
Mathematics	57%
Latin	57%
History	60%
	College students
English	35%
Mathematics	45%
History and economics	56%
Natural sciences	50%
Modern languages	40%

The difficulty with many of the measurements is that they are based on too small a number of persons. Comparisons based on twenty-five persons from either sex may be indicative but not final. Summarizing, we may say that women and girls are superior in sensibility, in memory, in most forms of perception, in quality of handwriting, and linguistic fluency. It is interesting to note in this connection that in the survey of mental-test results[1] the women excel in twelve out of fourteen tests which depend chiefly upon linguistic fluency. Thus the females excel in speed of reading, both oral and silent, in amount of information given in describing an object or in making a report, in the genus-species test, in the number of words thought of and written per minute, in the part-whole test, in the opposites test, in memory span for words, in memory for logical-verbal material, in the word-

[1] Given in the various chapters of Whipple's *Manual of Mental and Physical Tests.*

building test and in the Ebbinghaus completion test; while the males excel in the rate of association and in the sentence building test. Apparently the popular belief in the greater linguistic fluency of women is not without foundation. Men and boys are superior in motor capacities, such as tapping, quickness of reaction, in arithmetical reasoning, and in resistance to suggestions as indicated by the size-weight illusion and the use of suggestive questions in testimony. The two sexes seem to be approximately equal in associative processes and in most school subjects. The amounts of difference, however, are very small. This is particularly true of all the traits that have been measured in a sufficiently

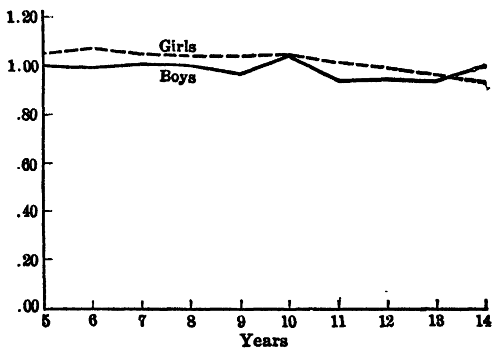

FIG. 29.—Comparison of general intelligence of boys and girls as measured by the Stanford revision of the Binet-Simon tests. After Terman ('16, p. 72). The numbers along the vertical axis are intelligence quotients as explained in Chapter VII.

large number of persons to make the comparisons safe. Any differences lying between 40% and 60% of the number of either sex reaching or exceeding the median of the other are practically negligible. If 60% of one sex reach or exceed the median of the other, it means that 10 persons in a hundred of the one sex, are by a small amount superior to the other. Differences larger than this have been established with a fair degree of certainty practically only in the case of one large field of capacities, namely, that of motor abilities. Differences in nearly all other respects in which comparisons have been made on large numbers of persons are almost entirely within the limits of 40% to 60%. Terman found in measuring the general intelligence of nearly 1,000 boys and girls by means of his revision of the Binet-Simon tests that for the ages of five to fourteen girls tend to be very slightly superior to boys and that after fourteen they are practically equal. His results are set forth in Figure 28.

It seems a likely interpretation that motor superiority has been carried over to include intellectual superiority as well. For centuries women have been considered intellectually inferior to men. They were thought to be incapable of acquiring anything more than an elementary education. It has been only since the middle of the 19th century that co-education and women's colleges have been generally established. Intellectual inferiority has probably

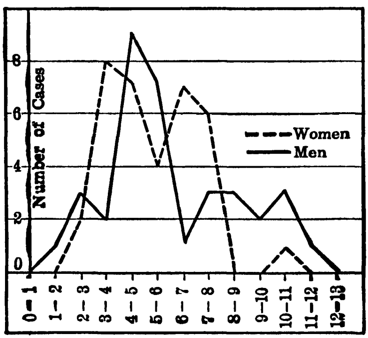

FIG. 30. Range of ability of men and women in color discrimination. After Henmon ('10).

been inferred chiefly from motor and muscular inferiority and from the conditions of a narrower environment and dependency due to the bearing and rearing of children. The inference and belief of intellectual inferiority is apparently unfounded. This conclusion may be fairly drawn both from the specific psychological tests that have been cited and also from the recent successes of women in the acquisition of higher education.

Difference in the Range of Variations in Abilities. Besides comparing the average amounts of any given ability in the two sexes, we may compare also the range of abilities from the lowest to the highest in the two sexes. Such comparisons have been made in a few traits and the general inference has been that the range of abilities is wider among men than among women. The distribution of the abilities in the geometrical perception test made upon 193 men and 200 women mentioned in a preceding paragraph, was as follows:

Scores:	2–3	4–5	6–7	8–9	10–11	12–13	14–15
193 Men	4.5%	15.4%	33.9%	21.9%	15.4%	5 6%	2.8%
200 Women	3.2%	20.8%	38.9%	21.6%	10.7%	3.9%	0.8%

Thus in the extremely high ability of canceling 14 to 15 geometrical figures in one minute, there were 2% more men than women, and in the lowest ability of canceling only two to three geometrical figures, there were 1.3% more men than women. Comparisons of this sort can be made safely only on large numbers of individuals, and consequently there is as yet little reliable material available.

The ratio of female to male variability has been computed by Thorndike ('14, III, p. 194) on the basis of tests of memory, reaction-time, discrimination of length, opposites, and cancellation made by Gilbert ('94) upon 100 boys and 100 girls of each age from 6 to 17. The average ratio in all tests for the ages of 9 to 12 was found to be .92, for the ages 13 to 14 1.025, and for the age of 15, .97. Girls were slightly less variable at all ages except 13 and 14. In a test of color discrimination Henmon ('10) also found a slightly larger variability among men than among women as shown in Figure 30.

The author made a comparison of the range of abilities in history and geography as measured by his tests in these subjects, and found the following distributions:

History, 8th Grade
Percentages of boys and girls attaining the various scores

Scores:	0–10	11–20	21–30	31–40	41–50	51–60	61–70
Boys	4.2%	9.3%	15.3%	17 0%	13.2%	12.2%	11.5%
Girls	6.2%	22.7%	22.7%	16.4%	12.8%	9.4%	6.0%

Scores (continued):	71–80	81–90	91–100	100–110	Total
Boys (continued)	9.3%	6.6%	2.1%	0 4%	288
Girls (continued)	2.3%	2.6%	0.9%	0.0%	352

Geography, 7th Grade

Scores:	0–10	11–20	21–30	31–40	41–50	51–60	61–70	71–80
Boys	06%	2.8%	4.7%	8.4%	5.3%	12.2%	14.7%	14.7%
Girls	03%	1 6%	5 6%	10.0%	10.0%	12.5%	18.7%	18.7%

Scores (continued):	81–90	91–100	101–110	111–120	121–130	Total
Boys (continued)	10.3%	8.1%	9.1%	6.9%	2.8%	320
Girls (continued)	10.0%	8.7%	7.2%	5.9%	3.1%	322

The variability of boys in the case of history is somewhat larger than that of the girls, whereas in the case of geography it is substantially the same.

Thorndike ('14, III, p. 195) has given the range of ages of boys and girls in the third year of high schools in Chicago, Philadelphia, New York, Detroit, Fall River, Los Angeles, Lowell, and Worcester as follows:

Age	13	14	15	16	17	18	19	20 and over	Total
Boys	7	92	594	1246	1203	572	193	67	3974
Girls	4	73	562	1351	1289	554	120	34	3987

There are about twice as many boys as girls at either 13 or 20 or over.

In support of the general belief that the range of general abilities is wider in men than in women may also be cited the fact that in the history of the world most of the great geniuses have been men, and also the statistical fact that male idiots and criminals at the other extreme of the distribution curve considerably outnumber the female. The fact that the great geniuses of the world have been men rather than women would accordingly be explained, not on the basis of lack of opportunity, but mainly on the basis of greater exceptional ability. The theory seems plausible but has been proposed rather in advance of a convincingly wide range of experimental data. If it is true, it would mean that according to the perception test the one or two per cent most gifted individuals are men and the 1 or 2% least gifted individuals are also men, that of the next 10 or 12% of most gifted individuals approximately two-thirds would be men and one-third women, and likewise of the next 10 or 12% least gifted individuals at the other extreme, about two-thirds would be men and one-third women. For the remainder of the distribution the number would be practically identical. The facts should not be interpreted as implying that men as a rule are superior to women, but would mean simply that only the one or two exceptional persons in a hundred would be superior to the most gifted women. The remaining 96 or 98% would be largely identical.

CHAPTER VI

THE INHERITANCE OF MENTAL TRAITS

Problem. In a certain obvious sense, the entire native equipment of any human being is inherited. The various capacities and the relative amounts of them with which a person starts in life are derived from the cells from which the individual originates. The differences among these original cells, even when derived from the same parent, are assumed to vary with regard to any potentiality according to the normal distribution curve about the central tendency of that particular parent. Stalks of corn grown from seed taken from the same ear will vary considerably from one another because the seeds themselves, even from the same ear, are different, but yet they will vary around the general type of the parent stalk. It is therefore obvious that children of the same parents will not be absolutely alike but that they will vary about the central tendency of their ancestors. The specific problem is not: Are mental traits inherited? but rather: How much do children of the same parents or ancestors resemble one another in the amounts of different traits possessed, and in the manner in which the various traits combine? To what extent are abilities in school work inherited? To what extent are the wide ranges of abilities, noticed in Chapter III, due to native equipment or to opportunity and environment? To what extent does a person make of himself what he does by virtue of his opportunities or by virtue of his inherent make-up? What part of the future adult individual is really determined by the school as an agency of his environment and what part is beyond the control of the school?

Methods of Studying Heredity. Any individual is the resultant of the interplay between his inherited equipment and the stimuli from his environment. Hence, theoretically, there are two general methods of studying the problem: First, by keeping the environment constant and varying the ancestry, so to speak; or second, by keeping the ancestry constant and varying the environment. That is, according to the former plan we would place children of entirely different ancestry into the same environment from birth up to a given point in life, and then measure the amount of similarity or difference; or according to the latter plan, we would place children of the same ancestry into entirely different environments from birth to a given point in life, and then measure the amount of similarity or difference. Such ideally scientific conditions are practically impossible to obtain. The best we can do is to measure the resemblances or differences of children of the same ancestry and compare them with the resemblances or differences of children of different ancestry, both groups living in approximately the same environment.

General Views Concerning Mental Heredity. Two extreme views concerning heredity are possible according to our conception of the relative rôles played by heredity and environment in the production of adult individuals. We may assume on the one hand that what a person becomes is absolutely and entirely determined by heredity, and that environment makes no difference whatever; or we may assume on the other hand that what a person becomes is completely and entirely determined by his environment, and that heredity plays no part. Neither view has been held by any serious student of heredity in recent times. Views very closely approaching these extremes, have, however, been held by prominent writers and thinkers in times past; whereas various views between these extremes are generally being held at the present time, depending upon the conception as to whether the larger, smaller, or equal share is contributed by heredity or by environment. The view held by most scientific students of the problem to-day gives weight to both elements with perhaps the major emphasis upon heredity.

The Similarity of Abilities among Related Eminent Persons. This particular method of attacking the problem was historically the first means of approaching the study of the inheritance of mental traits. Two extensive investigations on this aspect of the subject have been made. The first was carried out by Sir Francis Galton and published in 1869. Galton made a study of 977 eminent men, each of whom was the most eminent among 4,000 persons. He proceeded to determine how many relatives of equal eminence and of varying degrees of relationship each person possessed. In this manner he found that these 977 men had the following relatives of a like degree of eminence: 89 fathers, 114 brothers, 129 sons, 52 grandfathers, 37 grandsons, 53 uncles, and 61 nephews, or a total of 535. Galton further pointed out that 977 ordinary men selected by chance from the population at large would have only four such eminent relatives. He concluded as follows:

"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."

More recently an extensive study was made by Woods ('06) on mental and moral heredity in royalty. Woods made a comparison of 671 members of royal families in Europe by giving each person a rating on a scale of 1 to 10 in which 10 signified exceptionally high ability or genius, and 1 represented exceedingly low ability or imbecility. These ratings were made by the judgment of Woods himself according to the reports of these persons in histories and biographies. On the basis of these estimates, a tabulation was then made of the relationship of persons of various degrees of ability. He found that most of the eminent persons were grouped about four stocks or families out of fifteen, namely, the families of Frederick the Great, Queen Isabella of Spain, William the Silent, and Gustavus Adolphus. Likewise, he found that most of the persons of lowest ability were grouped around certain families in Spain and Russia, and the persons of mediocre ratings, four to seven, centered about some half dozen royal families including the houses of Hanover, Saxe-Coburg-Gotha, Reuss, Mecklenburg, Hapsburg in Austria, Holstein, Denmark, Saxony, Savoy, Orleans and modern Portugal. The ratings, of course, were not absolutely correct measurements of their abilities, but they, no doubt, represented greater validity than general impressions would. He further computed coefficients of resemblance in intellect and morals as follows:

I. In intellect:	
Offspring and father	.30
" " grandfather	.16
" " great-grandfather	.15
II. In morals:	
Offspring and father	.30
" " grandfather	.175

Dr. Woods then attempted to determine whether or not thefact of accession to the throne by virtue of birth gave an individual greater opportunity for eminence. This he states in the following manner:

"There is one peculiar way in which a little more than half of all 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 to 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 positions; 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 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 No. 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% 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.)"

"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%) 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.)

While heredity no doubt plays an important part in the production of intellect and character the part attributed to it by Woods that it explains "at least 90% of the intellectual side of character in every case" is hardly warranted either by the findings of other investigators or by the results of Woods himself. His correlation between father and offspring is only .30.

Similarities of Abilities Among Related Defective and Low Grade Persons. Quite a number of studies have been made in recent years concerning the frequency with which defective persons are either distantly or closely related. One of the first studies was that of the Jukes reported by R. G. Dugdale in 1877. Max Juke, born in 1720, was a shiftless truant, who married an equally worthless woman. Up to 1877 there had been five generations with approximately 1,200 descendants among whom have been traced the following types of persons: 310 paupers, 7 murderers, 60 habitual thieves, 50 prostitutes, 130 convicted of crime, 300 died in infancy, 440 physical wrecks from debauchery, only 20 learned a trade, and 10 of these learned it in prison. The estimated cost to the State of New York has been put at approximately $1,000 a person. In contrast with this lineage, a comparison has been suggested with the Jonathan Edwards family, which had approximately 1,400 descendants in the same period of time. Among them there have been 120 graduates of Yale alone, 14 college presidents, over 100 professors, 135 books of merit have been written by various members of the family, and 118 journals have been edited by them. Aaron Burr was the only black sheep among them and he can certainly not be classed as an intellectually defective person. (Winship '00).

Poellman of Bonn (Guyer '16, p. 271) made a study of a family called the Zeros in which 800 descendants were traced through six generations back to a female drunkard. Among them were found 102 professional beggars, 107 illegitimate offspring, 181 prostitutes, 54 inmates of almshouses, 76 convicted of crime, and 7 murderers. The cost to the state was placed at $1,206,000.

More recently a very interesting study was conducted by Dr. Goddard of the Training School at Vineland, New Jersey. Dr. Goddard ('12) traced the ancestry of a young girl who had been brought to his institution. It was found that the lineage went back to a man, Martin Kallikak, a soldier in the Revolutionary War, who was the progenitor of two lines of descendants. (See Figure 31.) He had an illegitimate son whose mother was feebleminded. This was the establishment of line—A—which had, down to the time of the study, 480 direct descendants among whom were found the following: 143 feeble-minded, 292 unknown, 36 illegitimates, 33 prostitutes, 24 alcoholics, 3 epileptics, 82 died in infancy, 3 criminals, 8 keepers of disreputable houses, and only 46 normal individuals. Apparently human nature does not gather grapes of thorns or figs of thistles. After his return from the war, Martin married a woman of normal intelligence and from this lineage—B—there had come during the same period of time, 496 direct descendants, of whom all were normal individuals with the

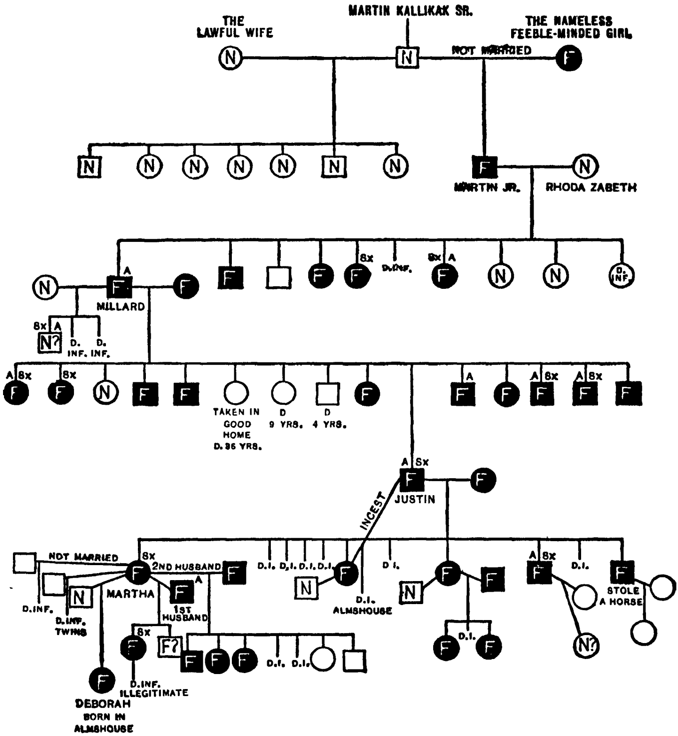

FIG. 31.—Descendants of the Kallikak Family. Squares = males, circles = females, black squares or circles = feebleminded, open squares or circles = normal persons. The lineage was traced back from Deborah. After Goddard.

exception of five, one of whom was reported as mentally defective, two as alcoholics, one as sexually immoral, and one as a case of religious mania. There were no epileptics or criminals, and only 15 died in infancy. The remainder were good citizens, including doctors, lawyers, educators, judges, and business men.

One thing seems to stand out very conspicuously from the numerous facts of family histories that have been unravelled in recent years, namely, that much defective mentality, degeneracy, and crime is a matter of ancestry. General opinion among persons in charge of institutions for defectives is that two-thirds of all cases are due to heredity and one-third to environmental or unknown causes. Thus Dr. Alfred Wilmarth, Superintendent of the Wisconsin Home for Feeble-minded, says:

"My own observations, and those of others in this country and Europe, would indicate that at least two-thirds of the feeble-minded have defective relatives. This is significant. Mental accident may occur in any family, but it is rarely a second case occurs unless there is a tendency to nerve degeneracy. (Quoted by Guyer,'16, p. 245.)

"I present to you the results of compiling the histories on 1,000 applications, where our information is most thorough; but I am confident that these do not tell the whole story. In 311 of these any neurotic taint in the family history is absolutely denied. In 365 cases at least one near relative suffers from one of the graver forms of nervous or mental trouble; in 170, two relatives were found; in 73 cases, three relatives, and in 81 cases four or more. These figures agree very accurately with the results of other observers in this country and abroad. It is safe to say that less than one-third of the defective classes are the results of disease or traumatism in families capable of transmitting a healthy, well developed nervous system."

Dr. Goddard of the Training School, Vineland, New Jersey, states in connection with his tests of 2,000 children:

"But we now know that 65% of these children have inherited the condition, and that if they grow up and marry they will transmit the same condition to their offspring. Indeed, we know that this class of people is increasing at an enormous rate in every community and unless we do something to stop this great stream of bad protoplasm we shall some day be swamped in a sea of degeneracy."

Likewise Dr. A. C. Rogers of the Minnesota School for Feeble-minded, at Faribault, says:

"We have no survey of mentality in this country except in very small areas, but probably about 65% of the feeble-minded children that we know of are feeble-minded from heredity; that is, they come from families in which there is much feeble-mindedness, usually associated with various neuroses or psychoses. There are about 35% approximately that are acquired cases. These cases develop from various things. Full development may be prevented during gestation, or early childhood, or early adolescence, but these acquired cases are entirely distinct from the hereditary ones." (Guyer '16, p. 246.)

Likewise, Dr. Martin W. Barr of the Pennsylvania Training School for Feeble-minded Children states:

"In my individual study of 4,050 cases of imbecility, I find 2,631 or 65.34%, caused by malign heredities; and of these 1,030, or 25.43%, are due to direct inheritance of idiocy; and 280, or 6.91%, to insanity."

To one who wishes to argue in favor of environment as the chief determining element in ability and character, such data as have been presented from family histories and relationships are not entirely convincing. It might be argued that a given family has so many individuals of high or low intelligence and achievement because its members were born in circumstances which did or did not afford opportunities for development and training and for achieving higher success. It might be said that the descendants of the Edwards family were born and reared among favorable circumstances of educational and financial advantages and consequently were fitted for greater tasks and lived in an environment in which larger opportunities offered themselves, whereas the members of such a lineage as the Jukes family would have just the opposite environment of birth, education, and opportunity in life. In answer to all this, we must remember, however, that ability very largely determines the sort of environment in which a person is satisfied to live, that a really capable person is quite likely to push forward and to find a way out of the environment in which he may happen to have been born, or to improve it if he cannot leave it, and finally, we must remember that the persons of low ability were born in circumstances of a correspondingly low nature because of the hereditary stock of the families from which they came. Their parents were content to live under the circumstances under which they did live because their abilities and desires sought for nothing better.

Similarities between Brothers and Sisters in Special Mental Traits. Pearson has shown that the resemblance in physical characteristics among brothers and sisters is approximately .50. He gives the following coefficients of correlation for various physical traits.

	Brother and sister
Hair color	.55
Cephalic index	49
Height	.50
Eye color	.52

What, however, is the degree of resemblance in mental traits? The general argument is that mental traits are dependent upon anatomical and neurological structures, and hence, if these are inherited, mental traits must also be inherited. What is the evidence from experimental and statistical facts?

In a study made by the writer ('17) a series of tests of capacities directly affected by school work and another series of tests of capacities not directly affected by school work were applied to 18 pairs of brothers and sisters in the University of Wisconsin. Each test was given twice on two different occasions in order to obtain a fairly accurate measurement of the capacities concerned. The purpose of giving the two types of tests was to ascertain whether brothers and sisters were more alike in the traits affected by training than in the traits not directly affected by school training. The following were the tests and the correlations obtained between pairs of children of the same family:

TABLE 18

Correlations between abilities of brothers and sisters. After Starch ('17)

Reading—speed	.51
Reading—comprehension	.64
Writing—speed	.72
Writing—quality	.46
Size of reading vocabulary	.07
Spelling	.05
Arithmetical reasoning	.38
Addition attempts	.71
Addition—rights	.44
Subtraction—attempts	.43
Subtraction—rights	.29
Multiplication—attempts	.37
Multiplication—rights	.25
Division—attempts	.46
Division—rights	.56
Average	.42
Memory	.31
A-test	.50
Geometrical form test	.07
Tapping	.65
Average	.38
Coefficients based on ranks in all tests combined	**.72**

In order to grasp the full import of these figures, it is necessary to remember that the coefficient of correlation between mental abilities of pairs of unrelated individuals selected by chance would be zero, and that any coefficient above zero between pairs of brothers and sisters means a corresponding amount of resemblance.

"From the above table several interesting results appear. (1) The resemblance of siblings is apparently no greater in those mental traits which are directly affected by school work than in those which are not so affected. The average correlation in the former group of tests is .42 and in the latter .38. This seems to indicate that the mental similarities of children of the same parents are due primarily to heredity rather than to similarity of environment since the resemblance is no greater in those traits which are more directly affected by environment.

"(2) The resemblance of siblings is approximately as great in mental traits as in physical traits. Pearson found the correlation between brother and brother in height to be .50 and in cephalic index (ratio of length to width of head) .49. These correlations for physical traits are a little larger than the ones found here for mental traits taken separately. The correlation, however, calculated on the basis of a combined rank for each person in all mental tests together was found to be .73. This greater correlation for all tests combined as compared with the correlation for single traits is due partly to the variation of the correlations among the single traits and partly due to the imperfections in the separate tests, which are counterbalanced to some extent in a combined ranking." (Starch '17, p. 237.)

Pearson ('04) made a comparison of 2,000 brothers and sisters who were rated by their teachers in such traits as vivacity, self-assertion, introspection, popularity, conscientiousness, temper, ability, and handwriting. On the basis of these ratings he found coefficients of correlation ranging from .43 to .64 with an average of .52. These results are interesting and indicative of the resemblance of more general traits of character, but they are probably rendered more or less uncertain by the unreliability of one person's ratings of such elements of character. The likelihood is that the teachers would be more inclined to estimate alike the children from the same families, rather than to estimate them more different than they really were.

Similarities of Brothers and Sisters in Abilities in School Subjects. In a study made several years ago, Earle ('03) found a correlation of .50 between the spelling abilities of 180 pairs of brothers and sisters. The writer ('15) made a study of the scholastic records of children from 63 families. The average grade in all school subjects was obtained for each pupil and used as the measure of his academic ability. The correlations based upon these averages were as follows:

First and second child in a family, 63 pairs	.58
Second and third child in a family, 24 pairs	.64
First and third child in a family, 24 pairs	.34
Average	.52

Further comparisons were made for abilities in specific school subjects which yielded the following correlations:

Spelling, 57 pairs of children from the same parents	.21
Reading, 57 pairs of children from the same parents	.49
Writing (speed) 24 pairs of children from the same parents	.18
Writing (quality) 24 pairs of children from the same parents	.06

Another study was made of 38 children from 11 families. All the marks that each pupil had received in each study in grades three to eight were averaged. From these averages the following coefficients of correlation were obtained.

Arithmetic, 54 pairs	.32
Spelling, 54 pairs	.21
Reading, 54 pairs	.31
Language, 54 pairs	.24

"No importance, I believe, can be attached to the differences in correlation between the various studies. The correlation for individual studies is lower than that for scholarship in general based on the average performance in all studies combined. This is probably due chiefly to the fact that the inaccuracies of teachers' marks in individual subjects are partly eliminated in the averages derived from all studies.

"Abilities in special subjects are inherited, apparently, to no greater extent in one subject than in another. What is probably inherited is either general scholarship or else more specialized traits than ability in arithmetic, or ability in language. Each study involves many mental faculties and nearly all studies involve the same faculties with varying emphasis.

"In corroboration of this point we may notice the following table of average marks for each of nine families in each study.

"In this table, we must examine the ranks, rather than the marks, of the different families in each subject, so as to eliminate the variation in standards of marking. These families rank very nearly the same in the various studies. For example, family C is first in every study except arithmetic and there it is third. Family G is second in every subject except arithmetic and there it is fourth. Family I is either third or fourth in every subject but one, and family B is last in every subject except one.

TABLE 19

Averages for each family in each subject.

Family	No. of Children	Arithmetic		Spelling		Reading		Language		Geography		History	
		Grade	Rank	Grade	Rank	Grade	Rank	Grade	Rank	Grade	Rank	Grade	Rank
A	3	72.3	9	85 1	5	82 2	3	80.7	5	76.7	8	77.9	8
B	5	75 1	7	76 1	9	73 2	9	74.6	9	75.3	9	76.1	9
C	4	80.2	3	89.1	1	86 7	1	83 6	1	84.6	1	84.0	1
D	6	80.4	2	85 9	3	81.1	5	81.6	4	78.6	6	78.5	6
E	2	77.4	5	78 5	7	74.4	8	76.1	7	79 3	4	78 7	5
F	4	73.6	8	80 8	6	76.1	7	76 0	8	78.6	6	78.2	7
G	3	77.6	4	86.6	2	84.4	2	83.2	2	83.8	2	83.8	2
H	3	76.1	4	78.4	8	80.5	6	80.0	6	78.6	6	79.6	4
I	3	81.8	1	85 5	4	82 1	4	82 5	3	81.6	3	81.9	3

"There is no evidence, at least from these figures, for the notion that special abilities in certain studies run in families. Mental traits running in families are very likely more specialized than abilities in school studies which involve large groups of mental functions. The children of any given family are on the average equally good or equally poor in all studies. Ability in school work is apparently inherited to the same extent as physical features since the coefficients of correlation for children of the same parents are approximately the same for both physical and mental traits." (Starch, '15, pp. 609–610.)

Schuster and Elderton ('07) calculated the resemblance in scholarship between brother and brother and between father and son among the Oxford honor men and found a coefficient of correlation of .40 for the former and .31 for the latter. Miss Elderton further determined the correlations between cousins from records of about 300 families and found a coefficient of .27.

Miss Emily S. Dexter [1] made a study of the scholarship records of 185 pairs of brothers, sisters, brothers and sisters, graduates of the University of Wisconsin, and of 69 similar pairs who were graduates of the high school at Ashland, Wisconsin. She reports the following coefficients:

[1] The study was carried out under the direction of Professor Henmon and reported in a thesis in the library of the University of Wisconsin, 1915.

TABLE 20

	Number of Pairs	General Scholarship in all Subjects	English	Language	Mathematics	History	Science
University:							
All pairs.......	185	.69	.64	.63	.55	.62	.60
Bro. and bro...	44	.47					
Sis. and sis....	71	53					
Bro. and sis....	66	62					
High School:							
All pairs......	69	.64	.58			.63	.61
Bro. and bro...	10	.38					
Sis. and sis....	26	39					
Bro. and sis....	23	36					

Miss Dexter concludes "that inheritance, to a much greater extent than training is responsible for the degree of resemblance found."

"If it were largely training, we would expect to find the resemblance greater between brother and brother, and sister and sister, than between brother and sister, but such is not the case. In the high school the correlations for the three groups are much the same, but, as has been pointed out, that may be due to a great extent to chance, for the groups are small. However, in the case of the university, where the groups average nearly three times as large as in the other school, we find the resemblance between brother and sister to be greater than between brother and brother, or sister and sister.

"Again, there is the question as to specialized abilities, and also that of general mental ability rather than specialized abilities as a basis of explanation for the close resemblance found. Thorndike, as has been said, finds that heredity is highly specialized. This study, however, seems to show a stronger tendency toward general mental ability, if by that we mean approximately equal ability in all subjects. It seems, also, to give almost no evidence of alternate inheritance; that is, of one individual's inheriting ability in one line while his brother inherits ability in another. In other words, a student who is above the average, either of his family or of the school, in one subject, is usually also above in most, and in many cases all, other subjects."

Similarities of Twins in Special Mental Traits. The two principal investigations on this phase of mental heredity were made by Galton and Thorndike. Galton made a general comparison of two groups of twins, one group of 35 pairs, which were reported as being very similar, in fact so similar that they were frequently reported as indistinguishable, and another group of twenty pairs of distinctly dissimilar twins. The conclusion formulated was to the effect that the former twins remained very similar all through life in spite of different environments, while the latter twins remained different all through life in spite of similar environments. Concerning certain of the twins Galton reports:

"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.'

"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.'

"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.'" ('83, p. 170, Everyman's Library Edition.)

Galton's general impression of his results is as follows:

"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." ('83, pp. 168 and 172.)

Thorndike's investigation ('05) was made by more accurate methods. He applied the tests, mentioned in the following table, to 50 pairs of twins and found the following correlations:

In the A-test	R—.69
In the a-t and r-e test	R—.71
In the misspelled word test	R— 80
In addition	R—.75
In multiplication	R— 84
In the opposites test	R— 90

"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 environments, then (1) twins should, to the age of leaving home, grow more and 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 in age should show a resemblance somewhat nearly as great as twin pairs, for the home and school conditions 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 a 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 equalling 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 equalling 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: (Thorndike '14, III, pp. 248–249).

TABLE 21

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 "

The Influence of Uniform Environment Upon Different Original Abilities. All studies cited thus far have attempted to measure the amount of similarity in related persons as compared with unrelated individuals on the assumption that the environment was roughly constant for all, that whatever resemblances existed between pairs of brothers and sisters or between other types of relatives greater than that between any pairs of persons selected by chance from the population at large, is considered to represent the actual amount of similarity in the original inherited natures of the individuals springing from the same ancestry. The problem may, however, be pursued further from a different angle, namely, by specific control of the environment or some portion of it. Thus it is possible to keep some particular part of the environment uniform for a group of persons of widely different abilities and to measure to what extent the original differences remain constant, increase, or decrease. If the differences remain constant, or increase, the inference would be that the ultimate differences of achievement would be primarily due to the inherited differences of capacities. If the differences decrease materially and finally disappear, the original differences would be mainly due to the effect of environment and opportunity.

A number of such experiments have been carried out. An investigation made by the writer ('11) in which 8 persons multiplied mentally 50 3-place numbers by a 1-place number each day for 14 successive days showed the following amounts of improvement:

TABLE 22

	Number of Examples in 1st 10 Min.	Number of Examples in Last 10 Min.	Gain in No. of Examples	Per Cent Gain
Three best persons ..	39	84	45	115
Three poorest persons	25	51	26	104

Hence, both the greatest absolute and the greatest relative gain was made by the group with the highest initial records.

Similar results have been found in the practice experiments of substituting numbers for letters as described in the author's *Experiments in Educational Psychology*, Chapter X. The following table gives the highest five and the lowest five records from among twenty persons. Each person practiced 120 minutes.

TABLE 23

Average number of letters transcribed.

	First 5 Min	Last 5 Min	Gain
Initial highest five persons..........	139	310	171
Initial lowest five persons..........	100	239	139

Again the largest gain was made by the group having the greatest initial ability.

Results pointing in the same direction have been obtained by Thorndike, Whitley, and others. For example, Thorndike ('10) found in the case of practice of nineteen persons in adding, the following results:

TABLE 24

The effect of equal amounts of practice upon individual differences in column addition of one-place numbers. After Thorndike ('10).

	Average Number of Additions 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

The statistical studies of scholastic histories of pupils through various periods of school life, which were discussed in a preceding chapter under the heading of correlations of abilities at different times of life in the same individual, all tend to corroborate the experimental facts here presented. The scholastic records show to a remarkable extent the uniformity with which each individual maintains his position throughout his educational career. A very interesting tabulation was made by L. J. Coubal and E. VanLandegend [1] to show the progress made by pupils in grades 4, 5, and 6 in one school in the four fundamental operations in arithmetic. Progress was measured by the Courtis tests, Series B, month by

[1] Under the direction of Professor Henmon, and reported in two theses in the library of the University of Wisconsin, 1917.

month through an entire year. The records in the four operations for each pupil were combined into a single score. Figures 32,

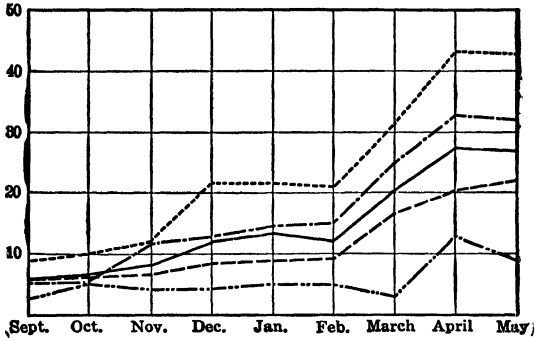

FIG. 32.—Progress in the four fundamental operations in arithmetic as measured by the Courtis tests, Series B, given at monthly intervals. The heavy continuous line represents all the pupils of the 4th grade. The four broken lines represent these pupils divided into quartiles.

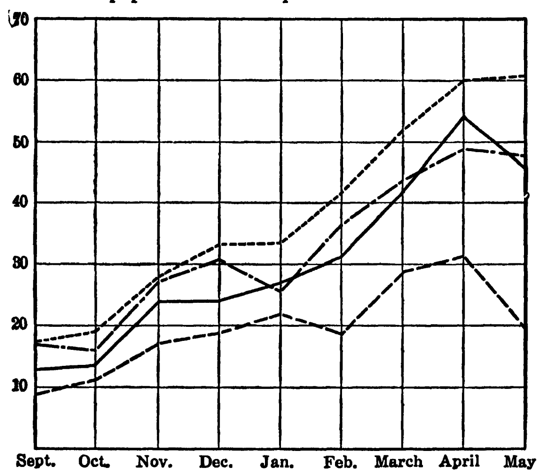

FIG. 33.—Same as Fig. 32, for the 5th grade. Pupils divided into three groups instead of four.

33, and 34 give the curves of progress for the respective grades. The pupils in each grade were divided into groups according to their final performance. Thus the pupils in grade 4 were divided into four groups while those in the other grades were divided into three groups. The results reveal the significant fact that the best groups in each grade made the greatest progress, the poorest groups made the least progress and the intermediate groups made average progress. The graphs for the various groups in any grade gradually spread apart during the course of the year, indicating that the differences increase rather than decrease or remain constant. The more gifted pupils profit more by their school work than the less gifted.

All experimental results point in the direction that practice does not equalize abilities; in fact, equal practice tends to increase differ-

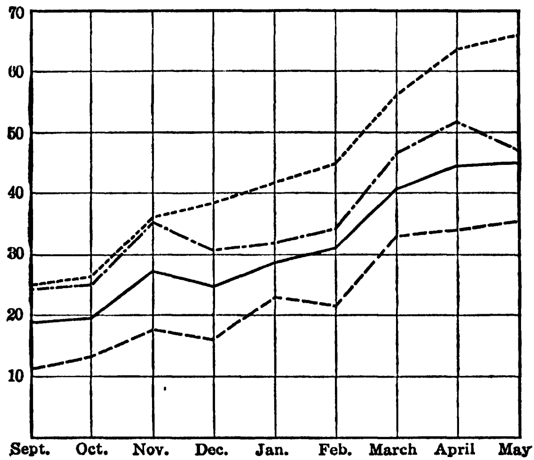

FIG. 34.—Same as Figs. 32 and 33, for 6th grade.

ences in achievement and skill rather than to decrease them. The more gifted individuals profit more, both relatively and absolutely, than the less gifted. This experimental fact is one of the most profound bits of evidence regarding the whole problem of heredity and environment. The talented men not only start with greater initial capacities but seem also to be capable of more intense application and more zealous desire to improve. "To him that hath shall be given" is psychologically true in the sphere of intellectual training as well as in the sphere of morality and religion. The man with ten intellectual talents will acquire far more than the man with one talent. If we may generalize for life as a whole, equal opportunities for all do not produce equal abilities in all. Men may be born free politically, but they are not born equal mentally; they may be born equal in opportunities in a democratic society, but they certainly are not equal in their ultimate achievements in life.

Influence of Different Environments upon Various Original Abilities. Extensive inquiries into the effects of various environmental conditions upon the native ability of human beings have been made in other fields besides the experimental one which has been surveyed. Such investigation as the study made by Dr. Rice ('97 and '02) upon the effects of various factors in school life upon the attainments in spelling and in arithmetic, the studies of places of birth of American men of science made by Cattell ('06), or the study of the places of birth of eminent men of letters made by Odin, and similar investigations by De Candolle ('73), Jacoby ('81), and Ellis, have been extensively referred to as bearing upon the problem of environmental forces in their interplay with hereditary capacities. The real significance and argumentative weight of such data seem to the writer to be uncertain and duplex in their meaning. Cattell, for example, has pointed out that the number of eminent scientific men born about 1860 in Massachusetts per one million population was very much greater than the number of eminent scientific men born in proportion to population in other states. To cite a few instances, he has computed that per one million population there were born eminent scientists as follows in various states:

Mass.	108 8
Conn.	86.9
R. I.	25.6
N. Y.	47 0
Wis.	45.0
Ill.	24.0
Ala.	2.0
Miss.	1.0

Similar figures are given for other states, and the inference made by Cattell is that the environment of Massachusetts and similar states has been much more conducive to the development of scientific men and that the number of such men could be determined practically by the control of the proper educational stimuli.

Odin, in his study of 5,233 noted French men of letters living during the period 1400 to 1830, found the following distribution according to places of birth:

1,229	born in	Paris
2,264	" "	other large cities
1,265	" "	small cities
93	" "	country districts

From this it has been inferred that if France as a whole had been as fertile as Paris in the production of eminent men of letters there would have been approximately 54,000 great men of letters instead of less than 6,000. The difficulty, however, with both studies is that Paris and Massachusetts have been more productive of eminent men not necessarily on account of better educational and social environment, but possibly also because of the fact that eminent men of letters and science have by virtue of the location in them of educational institutions, scientific and other intellectual centers, necessarily been attracted to these places, and consequently their children were born in these localities. The facts as such may actually be used in the support of heredity as against environment as much as they have been used in support of environment as against heredity.

Likewise, the study of Rice with regard to the factors affecting efficiency in school subjects is uncertain. Rice, on the basis of extensive tests in spelling and arithmetic in various parts of the country, arrived at the general opinion that practically all external conditions of home and school such as foreign or American parentage, home study, methods of teaching, size of class, and time devoted to study, made practically no difference whatever in the ultimate achievement of the pupils, and the implication is made that the final efficiency depends primarily upon heredity. The obvious uncertainty of such data as these is that while the facts in toto may imply such a situation, it is also quite certain that such a massing of data in this manner obliterates the effect of individual factors. Favorable conditions may be offset by unfavorable conditions and thus obscure the entire situation. To infer that good teaching and poor teaching make no difference in the ultimate results obtained, or that the amount of time given to study makes no difference in results, are conclusions that are quite likely to be unsound. The reason for the inference drawn by Rice is probably the fact that good teaching in some schools may be accompanied by other factors which tend to counteract its effect, whereas poor methods of teaching in other schools may be accompanied by favorable or unfavorable circumstances in other respects. The massing together of returns from many schools is bound to obliterate the effects of the individual conditions. The only certain way to ascertain the effectiveness of one factor or another would be to control all conditions, or to be able to allow for them definitely, with the exception of the one factor whose efficacy is to be determined. Thus in order to determine whether or not the different methods of teaching a given subject make a difference, it would be necessary to take a class of pupils in a given school and divide it up into two or more groups on the basis of equal initial capacities and to have each class taught, preferably by the same teacher and under the same general circumstances, by a different method. Then at stated intervals the two groups should be compared by reliable measures. In this manner definite results could be obtained as to the effectiveness of the various methods, conditions, or amounts of time devoted to a subject. Consequently, Rice's figures as they stand are of little worth so far as proving the forcefulness of different environmental factors in the production of results is concerned.

A similar criticism applies to such studies as that made by Spillman with regard to the place of birth of various men prominent in public and business life, such as presidents of the United States, governors, and railroad presidents. Spillman ('09) has pointed out, for example, that 23 of the first 25 presidents of the United States were born in the country, that 41 out of 45 governors, and 47 out of 62 cabinet officers were born in the country. It is unsafe to argue that because a large percentage of the presidents of the United States or other prominent persons were born and reared in rural districts, that country life is more productive of ability and ambition. To argue with any certainty on the basis of such facts, it is necessary to know the ancestral antecedents of the persons springing from various localities and sections of the population.

General Interpretation. The general impression from all experimental, statistical, and historical material thus far accumulated on the problems of mental heredity would seem to be somewhat as follows: Barring paupers, invalids, and those suffering from want of food and shelter due to conditions beyond their personal control, and referring to all others living in the same community at the same time, the ultimate achievement of any given individual is due to his original ability, probably to the extent of 60 to 90%, and to actual differences in opportunity or external circumstances only to the extent of 10 to 40%.

The facts of heredity bear down so heavily that the impression gained of the large part played by it leads one almost to a fatalistic philosophy. One is almost inclined to believe that persons become what they do largely on account of their hereditary capacities, and that they are not in the least responsible for their own outcome; that if a person is born with great capacities he will achieve high distinction, and if he is born with mediocre or slender capacities, he will not achieve anything beyond his limits no matter what he may do. While it is certainly true that no one may achieve a position higher than his original capacities will permit, it does not follow that a mechanical, fatalistic view needs to be taken. Nature predominates enormously over nurture only in the relative and not in the absolute sense. This distinction must always be borne in mind in studies of heredity. In fact, in the absolute sense, nurture predominates enormously over nature. A Newton born among Australian bushmen would no doubt have become a remarkable bushman, but never a world-renowned scientist. The necessary stimuli of environment must be at hand to train and develop original capacities. The difference between relative and absolute achievement may be illustrated in any of the experimental results concerning the effects of equal practice cited in a preceding section. The fact that all individuals improve by practice shows absolute gain in performance or skill. The fact that the gifted ones maintain their lead or even gain in their lead is relative achievement. Before practice, no child can write; after practice, all normal children can write with more or less excellence. This is absolute gain. Before practice, some children have greater original capacities for learning to write; after practice, these same children maintain the same superiority. This is relative gain. A Newton and an ordinary bushman born and reared among bushmen would probably be superior and ordinary bushman respectively. A Newton and an ordinary bushman born and reared in New York City at the beginning of the 20th century would probably become, respectively, the one a great scientific, professional or business man, and the other an ordinary person, able to get on, earn a living, and enjoy life within the ordinary limits. The original abilities of ancient civilized peoples were probably very little different from the original abilities of modern civilized peoples. The differences are probably due to the transformation of the environment which is constantly being brought about through the efforts of man. A Newton born in a modern civilized community would have greater and different stimuli than one born in an ancient or uncivilized community. His ultimate eminence would be determined by his environment.

The pessimistic air may further be dispelled by noting the fact that hardly one person in a thousand makes all the absolute gain possible for him even in a single capacity. It has been proved over and over again in numerous abilities which have been used daily in one's occupation that by a little special practice each day their efficiency may be enormously improved. Consequently, while the possibilities of each individual are limited by his original inherited equipment, each one may develop his capacities far beyond the usual degree of attainment. While experimental evidence indicates emphatically that under equal opportunities the more gifted surpass the less gifted, yet rarely does anyone do his best or attain his limit even in a single capacity. Life is a matter of competition; let everyone compete to the fullest extent of his inherited ability.

CHAPTER VII

THE MEASUREMENT OF MENTAL CAPACITIES

Problem. Strictly speaking, it is impossible to measure directly the original equipment of a human being unmodified by environmental causes. The nearest approach would be the preparation of a complete inventory, and an exact measurement, of all the capacities that an individual possesses at birth. Even then, pre-natal conditions have entered into the growth of the organism. The next nearest approach would be a measurement of all capacities which are not directly or specifically trained by school, occupation, or special circumstances. In fact, no one can live and possess capacities without any modification of them from outside causes; hence the best that we can do is to measure as many capacities and abilities as possible which have been modified least by special exercise or training, and then to consider them as approximately representing a person's original abilities, or, to make such allowances as we can for the influence of external causes. No human being up to the present time has been measured in all respects at any given point in his growth by thoroughly accurate methods. A great many persons, however, have been partially measured in a great many capacities by more or less accurate or inaccurate methods at various stages of their growth.

General Value. John Stuart Mill has said the "greatest thing in the world is man, and the greatest thing in man is mind." To this statement we might possibly add that the greatest achievement of science would be the measurement of the mind. The import and value of definite means for measuring the capacities of human beings would touch all phases of human life in which intelligence is involved. If we had accurate means of describing a given person's capacities in all directions in terms that could be precisely defined and understood, we would have an instrument for evaluating human beings far beyond our present possibilities. We would then be able to obtain a precise notion of the capacities of an individual. Consider for a moment what the advantages would be! In all sorts of human relations, men are called upon constantly to pass judgment upon their fellows concerning their fitness, capacity, and promise of success for this or that particular line of work. The one thing which is probably most important of all, aside from the special training in a given field, is the intelligence and native ability which a person possesses. What aptitudes does a person have for this or that type of work? Enormous waste in the energies of men are due to mal-adaptation of individual to task. The business world is rapidly turning toward psychology for help, and if psychology is to give the help it will have to be in the form of adequate measurements of the capacities of human beings. Sound vocational guidance, in which much interest has recently sprung up, will have to be founded upon a sound vocational psychology whose development lies largely in the future. Courts are recognizing that responsibility for conduct rests upon mental maturity and intelligence, and that these must be determined first before proper adjudication may be made of an individual's behavior. Psychological laboratories have therefore been established in recent years in connection with juvenile courts. The immigration office finds it necessary to make intelligence examinations, even if crude, in order to exclude those distinctly unfit. In normal times a considerable number, 80 to 100 per month, are returned to the countries whence they came on account of mental deficiency.

One of the large problems of the school is the proper adjustment of work and progress to the natural ability of the pupils and, in particular, the discovery of the morons and borderline individuals so that they may be taken care of in special classes or otherwise to the best advantage to themselves and to the other pupils in the school. Intelligence examinations would be useful not only in connection with the relatively small percentage of backward and defective pupils, but also in connection with the normal and superior individuals. Such tests would be valuable in conjunction with the measurement of attainment in school subjects specifically, so that a child's progress and rate of advancement could be determined on the basis of both types of measurements. Precise methods of evaluating the actual capacities of pupils would be of decided value in making possible a more accurate promotion or retardation of pupils according to their abilities, and a more accurate prescription of work to be done and of the progress that can most profitably be made. The school has paid relatively more attention to the backward pupils by putting them into special classes than to the superior ones. And yet the latter will be the ones who will contribute most to the advancement of society as a whole. Why should there not be special classes for the gifted pupils so that they might be led to reach their fullest intellectual growth and thus return to society the most that they are capable of?

Methods of Measuring Original Capacities. In general two types of methods have been developed, at least in part, and used for determining the native ability of human beings. The term "native" of course, must be understood to signify not pure, native ability unmodified by experience, but native or original only in the sense of not being directly affected by specific training. The one method consists of a considerable variety of reactions to questions and situations which a child would be able to make as a result of normal growth in a normal environment. The tests developed on this principle are the Binet-Simon tests and the various modifications of them.

The second general method has proceeded on the basis of measuring, by fairly precise methods, certain special mental functions from year to year, and of determining thereby the mental status and growth of the individual. Thus, for example, many capacities might be measured by a definite psychological technique from year to year, and certain norms might be established for each year so that we could say that a given individual's memory has been developed to the norm or average of a child of ten. Similar tests and norms could be developed in as many different mental capacities as would seem to be necessary in order to obtain a fairly complete evaluation of an individual's natural abilities. This second general method has not as yet been developed to the same degree of completeness as the Binet-Simon type ('05) with respect to either the selection of the particular capacities that should be tested, or the types of tests that ought to be used, or the technique by which they should be given. Brief consideration will be given to both plans of measurement.

The Binet-Simon Scale. This series of tests is arranged in groups according to years. Thus there is a series of tests for every year from age three up to twelve or fifteen, and in some of the revisions even to adult life. These tests were first prepared by the French psychologist, Binet, and the French physician, Simon, who collaborated for a period of twelve or fifteen years in the selection of tests and in assigning them to the proper years according to the growth and development of the child. These tests were first published in 1905 and since then were revised by the original authors in 1908 and in 1911. A number of investigators have attempted to revise them and to adapt them to the conditions in their respective countries, and to improve them so that they would be more reliable and more accurately graded according to the growth of children from year to year. In this country, the chief revisions and improvements have been made by Goddard, Kuhlmann, Yerkes, Terman, and others. Probably the most satisfactory and careful revision of the original Binet tests is the one recently prepared by Terman and known as the Stanford Revision. This revision consists in the elimination of some of the original tests, in the addition of a considerable number of new tests, in the readjustment of other tests up or down the scale of years according to their difficulty, and particularly in the development of a more precise technique for giving and evaluating the tests, so that examiners may be guided specifically in the administration of them. The following is a complete list of the tests in the Stanford revision. The detailed directions for giving and scoring the tests together with extensive results, are given in Terman's *The Measurement of Intelligence.*

The Stanford Revision and Extension

Year III. (6 tests, 2 months each.)

1. Points to parts of body. (3 of 4.)
   Nose; eyes; mouth; hair.
2. Names familiar objects. (3 of 5.)
   Key, penny, closed knife, watch, pencil.
3. Pictures, enumeration or better. (At least 3 objects enumerated in one picture.)
   (a) Dutch Home; (b) River Scene; (c) Post-Office.
4. Gives sex.
5. Gives last name.
6. Repeats 6 to 7 syllables. (1 of 3.)

A1. Repeats 3 digits. (1 success in 3 trials. Order correct.)

Year IV. (6 tests, 2 months each.)

1. Compares lines. (3 trials, no error.)
2. Discrimination of forms. (Kuhlmann.) (Not over 3 errors.)
3. Counts 4 pennies. (No errors.)
4. Copies square. (Pencil. 1 of 3.)
5. Comprehension, 1st degree. (2 of 3.) (Stanford addition.)
   "What must you do?" "When you are sleepy?" "Cold?" "Hungry?"
6. Repeats 4 digits. (1 of 3. Order correct.) (Stanford addition.)

A1. Repeats 12 to 13 syllables. (1 of 3 absolutely correct, or 2 with 1 error each.)

Year V. (6 tests, 2 months each.)
1. Comparison of weights. (2 of 3.)
   3–15; 15–3; 3–15.
2. Colors. (No error.)
   Red; yellow; blue; green.
3. Æsthetic comparison. (No error.)
4. Definitions, use or better. (4 of 6.)
   Chair; horse; fork; doll; pencil; table.
5. Patience, or divided rectangle. (2 of 3 trials. 1 minute each.)
6. Three commissions. (No error. Order correct.)

A1. Age.

Year VI. (6 tests, 2 months each.)
1. Right and left. (No error.)
   Right hand; left ear; right eye.
2. Mutilated pictures. (3 of 4 correct.)
3. Counts 13 pennies. (1 of 2 trials, without error.)
4. Comprehension, 2nd degree. (2 of 3.) "What's the thing for you to do?"
   (a) "If it is raining when you start to school?"
   (b) "If you find that your house is on fire?"
   (c) "If you are going some place and miss your car?"
5. Coins. (3 of 4.) Nickel; penny; quarter; dime.
6. Repeats 16 to 18 syllables. (1 of 3 absolutely correct, or two with 1 error each.)

A1. Morning or afternoon.

Year VII. (6 tests, 2 months each.)
1. Fingers. (No error.) Right; left; both.
2. Pictures, description, or better. (Over half of performance description.) Dutch Home; River Scene; Post-Office.
3. Repeats 5 digits. (1 of 3. Order correct.)
4. Ties bowknot. (Model shown. 1 minute.) (Stanford addition.)
5. Gives differences. (2 of 3.)
   (Fly and butterfly; stone and egg; wood and glass.)
6. Copies diamond. (Pen. 2 of 3.)

A1. 1. Names days of week. (Order correct. 2 of 3 checks correct.)

A1. 2. Repeats 3 digits backwards. (1 of 3.)

Year VIII. (6 tests, 2 months each.)
1. Ball and field. (Inferior plan or better.) (Stanford addition.)
2. Counts 20 to 1. (40 seconds. 1 error allowed.)
3. Comprehension, 3d degree. (2 of 3.) "What's the thing for you to do?"

(a) "When you have broken something which belongs to some-one else?"
(b) "When you are on your way to school and notice that you are in danger of being tardy?"
(c) "If a playmate hits you without meaning to do it?"

4. Gives similarities, two things. (2 of 4.) (Stanford addition.) Wood and coal; apple and peach; iron and silver; ship and automobile.
5. Definitions superior to use. (2 of 4.) Balloon; tiger; football; soldier.
6. Vocabulary, 20 words. (Stanford addition. For list of words used, see record booklet.)

A1. 1. First six coins. (No error.)
A1. 2. Dictation. ("See the little boy." Easily legible. Pen. 1 minute.)

Year IX. (6 tests, 2 months each.)

1. Date. (Allow error of 3 days in c, no error in a, b, or d.) (a) day of week; (b) month; (c) day of month; (d) year.
2. Weights. (3, 6, 9, 12, 15. Procedure not illustrated. 2 of 3.)
3. Makes change. (2 of 3. No coins, paper, or pencil.) 10–4; 15–12; 25–4.
4. Repeats 4 digits backwards. (1 of 3.) (Stanford addition.)
5. Three words. (2 of 3. Oral. 1 sentence or not over two co-ordinate clauses.) Boy, river, ball; work, money, men; desert, rivers, lakes.
6. Rhymes. (3 rhymes for two of three words. 1 minute for each part.) Day; mill; spring.

A1. 1. Months. (15 seconds and 1 error in naming. 2 checks of 3 correct.)
A1. 2. Stamps, gives total value. (Second trial if individual values are known.)

Year X. (6 tests, 2 months each.)

1. Vocabulary, 30 words. (Stanford addition.)
2. Absurdities. (4 of 5. Warn. Spontaneous correction allowed.) (Four of Binet's, one Stanford.)
3. Designs. (1 correct, 1 half correct. Expose 10 seconds.)
4. Reading and report. (8 memories. 35 seconds and 2 mistakes in reading.) (Binet's selection.)
5. Comprehension, 4th degree. (2 of 3. Question may be repeated.)
   (a) "What ought you to say when some one asks your opinion about a person you don't know very well?"
   (b) "What ought you to do before undertaking (beginning) something very important?"

(c) "Why should we judge a person more by his actions than by his words?"

6. Name 60 words. (Illustrate with clouds, dog, chair, happy.)

A1. 1. Repeats 6 digits. (1 of 2. Order correct.) (Stanford addition.)

A1. 2. Repeats 20 to 22 syllables. (1 of 3 correct, or 2 with 1 error each.)

A1. 3. Form board. (Healy-Fernald Puzzle A. 3 times in 5 min.)

Year XII. (8 tests, 3 months each.)

1. Vocabulary, 40 words. (Stanford addition.)
2. Abstract words. (3 of 5.)
   Pity; revenge; charity; envy; justice.
3. Ball and field. (Superior plan.) (Stanford addition.)
4. Dissected sentences. (2 of 3.) (1 minute each.)
5. Fables. (Score 4; i. e., two correct or the equivalent in halt credits.) (Stanford addition.)
   Hercules and Wagoner; Maid and Eggs; Fox and Crow; Farmer and Stork; Miller, Son, and Donkey.
6. Repeats 5 digits backwards. (1 of 3.) (Stanford addition.)
7. Pictures, interpretation. (3 of 4. "Explain this picture.")
   Dutch Home; River Scene; Post-Office; Colonial Home.
8. Gives similarities, three things. (3 of 5.) (Stanford addition.)
   Snake, cow, sparrow; book, teacher, newspaper; wool, cotton, leather; knife-blade, penny, piece of wire; rose, potato, tree.

Year XIV. (6 tests, 4 months each.)

1. Vocabulary, 50 words. (Stanford addition.)
2. Induction test. (Gets rule by 6th folding.) (Stanford addition.)
3. President and king. (Power; accession; tenure. 2 of 3.)
4. Problems of fact. (2 of 3.) (Binet's two and one Stanford addition.)
5. Arithmetical reasoning. (1 minute each. 2 of 3.) (Adapted from Bonser.)
6. Clock. (2 of 3. Error must not exceed 3 or 4 minutes.)
   6.22; 8.10; 2.46.

A1. Repeats 7 digits. (1 of 2. Order correct.)

Average Adult. (6 tests, 5 months each.)

1. Vocabulary, 65 words. (Stanford addition.)
2. Interpretation of fables. (Score 8.) (Stanford addition.)
3. Difference between abstract words. (3 real contrasts out of 4.)
   Laziness and idleness; evolution and revolution; poverty and misery; character and reputation.
4. Problem of enclosed boxes. (3 of 4.) (Stanford addition.)

5. Repeats 6 digits backwards. (1 of 3.) (Stanford addition.)
6. Code, writes "Come quickly." (2 errors. Omission of dot counts half error. Illustrate with "war" and "spy.") (From Healy and Fernald.)

A1. 1. Repeats 28 syllables. (1 of 2 absolutely correct.)

A1. 2. Comprehension of physical relations. (2 of 3.) (Stanford addition.)
Path of cannon ball; weight of fish in water; hitting distant mark.

"Superior Adult." (6 tests, 6 months.)

1. Vocabulary, 75 words. (Stanford addition.)
2. Binet's paper-cutting test. (Draws, folds, and locates holes.)
3. Repeats 8 digits. (1 of 3. Order correct.) (Stanford addition.)
4. Repeats thought of passage heard. (1 of 2.) (Binet's and Wissler's selections adapted.)
5. Repeats 7 digits backwards. (1 of 3.) (Stanford addition.)
6. Ingenuity test. (2 of 3. 5 minutes each.) (Stanford addition.)

The mental maturity or intelligence of a child is expressed in terms of a quotient which represents the relation between his mental age and his chronological age and is obtained by dividing the former by the latter. Thus, a child 10 years old with a mental age of 10 would have an intelligence quotient (IQ) of 1.00; a child 10 years old with a mental age of 11, would have an intelligence quotient of 1.10, or a child 10 years old with a mental age of 8 would have an intelligence quotient of .80. If the quotient is under 1.00 it means that the child is below the average and if it is above 1.00, it means that the child is above the average. Terman has suggested the following classification of intelligence quotients with their approximate meanings:

"IQ	Classification
Above 1.40	"Near" genius or genius.
1.20–1.40	Very superior intelligence.
1.10–1.20	Superior intelligence.
.90–1.10	Normal, or average, intelligence.
.80– .90	Dullness, rarely classifiable as feeble-mindedness.
.70– .80	Border-line deficiency, sometimes classifiable as dullness, often as feeble-mindedness.
Below .70	Definite feeble-mindedness.

"Of the feeble-minded, those between .50 and .70 IQ include most of the morons (high, middle, and low), those between .20 or

.25 and .50 are ordinarily to be classed as imbeciles, and those below .20 or .25 as idiots. According to this classification the adult idiot would range up to about 3-year intelligence as the limit, the adult imbecile would have a mental level between 3 and 7 years, and the adult moron would range from about 7-year to 11-year intelligence."

Terman ('17) has attempted to estimate the boyhood intelligence quotient of Sir Francis Galton from such records as are available of his youth, and believes it quite certainly to have been 2.00. For example, on the day before his fifth birthday he wrote the following letter to his sister, the statements of which have been corroborated by other general evidence:

"My dear Adele,

"I am 4 years old and I can read any English book. I can say all the Latin Substantives and Adjectives and active verbs besides 52 lines of Latin poetry. I can cast up any sum in addition and can multiply by 2, 3, 4, 5, 6, 7, 8, (9), 10, (11).

"I can also say the pence table. I read French a little and I know the clock.

"Francis Galton,
"February 15, 1827."

At the age of 10, young Galton wrote the following letter which represents maturity of judgment and intellectual interest worthy of a high school or college student:

"December 30, 1832.

"My Dearest Papa:

"It is now my pleasure to disclose the most ardent wishes of my heart, which are to extract out of my boundless wealth in compound, money sufficient to make this addition to my unequaled library.

The Hebrew Commonwealth by John..........	9
A Pastor's Advice..........................	2
Hornne's Commentaries on the Psalms..........	4
Paley's Evidence on Christianity...............	2
Jones Biblical Cyclopedia.....................	10
	27"

To illustrate concretely the manner of determining the mental age of a child, the following record of the examination of a boy 14 years and 11 months old is cited. The test numbers refer to the preceding list.

Year III. (6 tests, 2 months each.)
1. Passed.
2. "
3. "
4. "
5. "
6. "

Year IV. (6 tests, 2 months each.)
1. Passed.
2. "
3. "
4. "
5. "
6. Failed.

Year V. (6 tests, 2 months each.)
1. Passed.
2. "
3. Failed.
4. "
5. Passed.
6. "

Year VI. (6 tests, 2 months each.)
1. Passed.
2. Failed.
3. "
4. "
5. Passed.
6. Failed.
A1. Passed.

Year VII. (Failed.)

This boy passed all the tests of the third year and twelve additional tests scattered through the years IV, V, and VI, for each of which he receives two additional months of credit. Hence his mental age is three years plus 24 months or five years, and his intelligent quotient is .34. He falls into the class of imbeciles.

A different plan of evaluation has been prepared by Yerkes and

Bridges ('15) in the construction of their point scale from the materials contained in the original Binet tests. This modification has proved quite satisfactory in practice. Haines has constructed from the same material a similar point scale which is adapted for use with the blind. Pintner and Paterson ('17) have assembled and standardized a series of fifteen performance tests chiefly of the form-board type which are especially valuable for use with the deaf and the mute.

It must not be assumed from the apparent simplicity of the Binet tests that any novice can use them. On the contrary it requires considerable training and psychological insight to use them properly. Some persons by reason of lack of tact and sympathetic attitude are temperamentally unfitted for mental testing. It is desirable that at least a year of training in experimental psychology should precede practice work in giving the tests; and before the examiner may be confident of administering them

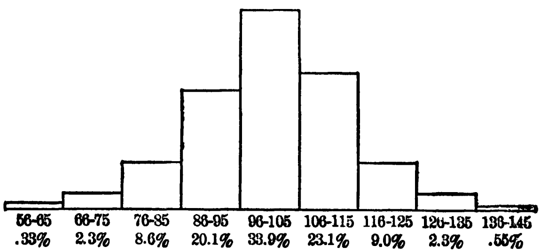

FIG. 35.—Distribution of intelligence quotients of 905 pupils. After Terman ('16, p. 66).

satisfactorily he should have practice under supervision with thirty to forty cases. The examining room should be plainly furnished and free from disturbing interruptions or distractions. The confidence of the child should be obtained before the examination is begun and his efforts should be encouraged so that he may react to his best advantage. Above all, the exact formula for giving and scoring each individual test must be followed rigorously if the tests are to have reliable diagnostic value.

Measurements Obtained from the Use of the Binet-Simon Tests. Goddard tested some 1,500 pupils by means of the revision of the Binet-Simon scale prepared by himself. He found the distribution of abilities as shown in the following table in which are specified the number of pupils one or more years ahead or behind their chronological age in mental maturity. The remarkable finding is the enormous range of the mental ages of pupils of the same chronological age, extending from those who are 4 or 5 years retarded to those who are 4 or 5 years accelerated.

TABLE 25

Chronological Age	5	4	3	2	1	At Age	1	2	3	4
4.........					1	2	2	3		
5.........			2	4	8	40	40	16	4	
6.........			1	0	29	48	69	9	0	1
7.........		1	2	8	15	114	50	4	3	
8.........		2	2	1	87	86	18	12	3	
9.........				27	54	56	56	4	2	
10.........			15	24	19	124	27	8	2	
11.........		4	13	25	50	60	12	1		
12.........	4	10	13	42	36	39	7			
Totals	4	17	48	131	299	569	281	57	14	1

A similar investigation was made by Terman ('16) of 905 pupils showing a distribution in terms of the intelligence quotient as given in the accompanying diagram, Figure 35. Kuhlmann ('12) tested some 1,300 defective individuals in the institution at Faribault, Minnesota, and found mental ages for various chronological ages as follows:

TABLE 26

Chronological Age	Number of Cases	Average Chronological Ages	Average Mental Age
1– 5	7	4.6	2.6
6–10	85	8.7	3.8
11–15	194	12.9	5.1
16–25	353	20.0	5.5
26–60	367	36.5	5.5

Many uses of such measurements as these may be made in school. Thus, for example, the number of retarded pupils in our schools is usually stated to be very large and the number of accelerated pupils is usually stated to be very small. The percentage of retarded pupils is placed in many schools anywhere from 25 to 50%. An intelligence examination of 483 pupils in Kansas City

(Dougherty '13), showed that the percentage of pupils as distributed both according to intelligence and according to over- and under-age plans was as follows:

	RETARDED	NORMAL	ADVANCED
Chronological Age.......	49%	45%	6%
Mental Age............	21%	49%	30%

It appears thus that the large percentage of so-called over-age pupils is not due to lack of capacity but to various other causes. The abilities of pupils seem to be distributed approximately in a normal, symmetrical manner. The large majority of pupils are in the center or at-age with about equal numbers retarded or accelerated. Many of the over-age pupils in the numerous age-grade statistics that have recently been complied are not mentally arrested but entirely normal or average.

Measurements of Special Capacities in Relation to General Intelligence. The second general method of measuring original capacities is as yet largely in the experimental stage. However, substantial beginnings have been made and a number of investigations have been carried out which indicate with considerable assurance that it would be possible to select a series of tests of specific mental functions which would be correlated closely with general ability and could, therefore, be used as symptomatic measures of general intelligence. An extensive investigation on this order was made by Simpson ('12) in which he attempted to solve two problems one of which was to determine what sort of tests would be indicative of intelligence, and the other was to apply them to two groups of persons widely separated in general intelligence. He proceeded accordingly to apply some 13 or 14 tests to two groups of persons, one group of 17 individuals composed of professors and graduate students of Columbia University, which represented the upper end of the intelligence scale, and the other group of 20 persons whom he found at the Salvation Army Industrial Home and at the mission on the Bowery in New York City, which represented the lower end of the intelligence scale. These tests were applied from two to four times each and were then compared with estimated intelligence as ranked by the impressions of several persons. The correlations of these various tests with estimated intelligence were as follows:

TABLE 27. After Simpson ('12)

Estimated intelligence and Ebbinghaus Test	.89
" " " Hard Opposites	.96
" " " Memory for Words	.93
" " " Easy Opposites	.82
" " " "A" Test	.21
" " " Memory for Passages	.35
" " " Adding	.72
" " " Geometric Forms	.07
" " " Learning Pairs	.34
" " " Completing Words	—
" " " Drawing Lines	—.20
" " " Extending Lines	—

From these correlations it appears that some of the tests serve very well as indicators of general ability, such as the various memory tests, the Ebbinghaus test, and the opposites tests. Certain other tests, however, seem not at all to be symptomatic of general capacity since their correlations are approximately zero, such as, for example, the drawing of lines and geometric forms.

A similar study was made somewhat earlier by Burt ('09), in which he found the following correlations of various functions with estimated intelligence:

TABLE 28. After Burt ('09)

General Intelligence and Attention (Dotting)	.72
" " " Apprehension (Pattern)	.75
" " " Adaptability (Mirror)	.60
" " " Memory (Words)	.67
" " " Discrimination (Alphabet)	.70
" " " " (Cards)	.54
" " " Reaction (Tapping)	.45
" " " " (Dealing)	.36
" " " Perception (Pitch)	.38
" " " " (Lines)	.23
" " " Touch (Two pt.)	.03
" " " Weight Discrimination	—.16

In an investigation made by the writer a series of eight tests was applied to a group of 15 high school pupils. Each test was given very carefully to each child individually according to a uniform technique and repeated four times on four different occasions. The difficulty with a great deal of testing work has been that the tests have not been applied under sufficiently constant or rigorous conditions or repeated sufficiently often to yield a fairly accurate measurement of the particular capacity in question. At the close of these tests each of the 15 pupils was asked to give a rating of the other 14, placing the one that was estimated to have highest intelligence as 1, and so on down to the 14th. The pupils knew one another well and so were able to give fair opinions. From these ratings a combined rank of general, estimated intelligence was obtained and these ranks were then correlated with the ranks in individual tests as well as with the combined ranks in all tests put together. The results are shown in the following table:

TABLE 29

Estimated Intelligence and Memory of Words	.52
" " " Memory of Passages	.56
" " " Opposites	.08
" " " Mental Addition	.50
" " " Arithmetical Reasoning	.83
" " " All Tests	.69
" " " " " except Opposites	.89

Thus it will appear that the tests individually, with the exception of the opposites, as well as collectively, agree very closely with the combined estimates of intelligence given by the 15 pupils of one another. The estimates of intelligence of pupils agreed very closely among themselves. The pupil who stood highest in the estimates of his fellows was estimated first by all pupils but one. The ranking of the others, of course, did not agree as closely, but the agreement was so close that the combined ranking yielded a rather reliable rating. The chief discrepancies between the estimated rank and the test rank occurred in the case of two of the fifteen pupils. The one was a boy who was estimated considerably lower by his fellows than was his rank in the tests. His estimated rank was 14 while his test rank was 5. The other pupil was a girl whose estimated rank was considerably higher, 5, than her test rank, 11. Upon inquiry it was discovered that the boy was not liked well by his associates, while the girl was unusually popular. Their true mental ability was probably indicated more correctly by the tests. The pupil who stood first in all tests combined, and was given first place by all his comrades except one, completed the high school at fourteen and maintained an excellent record in the university.

The promise of these and other investigations is sufficiently great to make it possible to develop a series of properly selected tests with a definite technique to measure the general ability of human beings with considerable trustworthiness. It would be necessary, after the selection of the tests, to develop definite norms for each test and for each particular year from infancy up to adult life. The advantage of a series of tests of this kind over the Binet type, would be that they could be applied and evaluated with greater precision; that they would measure more directly certain fundamental mental capacities; that they would be less dependent upon particular environmental conditions, and that they would yield more objective and scientific results. The specific test method of measuring intelligence gives greater scientific promise and will in the future probably replace the Binet method to a large extent.

PART II

THE PSYCHOLOGY OF LEARNING: A. IN GENERAL

CHAPTER VIII

ANALYSIS OF PROBLEMS

Analysis of the Learning Process. What are the mental and neural processes involved in various types of learning such as writing, reading, spelling, adding, solving problems, operating typewriters, grasping laws of nature, retaining facts, playing tennis, riding a bicycle, sawing to a straight line, speaking a foreign language, and the like? Probably all forms of learning can be reduced to one relatively simple, schematic type: Reception of impressions through the senses; assimilation, analysis, and combination of processes in the mind; and redirection of impulses to produce a reaction; or in brief, stimulus, association, response. A child learns to avoid a disagreeable stimulus by receiving the sensation, associating it with disagreeable consequences, and reacting by avoiding the stimulus. The chick learns to avoid disagreeable caterpillars by the same process. Visual and gustatory stimuli are brought to the appropriate centers of the brain and there associated so that whenever in the future the visual stimuli of the caterpillar are brought in, the disagreeable taste associations will also arise which will cause an inhibition of the muscular actions concerned in pecking at the caterpillar. A pupil, on the first day of school, is shown certain black marks on a chart or in a book and is told "hat" which is to cause him also to say the word "hat." The psychological series of events is as follows: First, the visual stimulus from the page transmits impulses to the visual centers in the brain and simultaneously the auditory stimulus from the pronunciation of the word by the teacher transmits impulses to the auditory center in the brain; second, connections between these visual and auditory stimuli in the brain and arousal by the auditory stimuli of images and meanings of the object "hat" which have been established through previous experiences before coming to school; and third, a redirection of impulses to the motor centers to attempt to speak the word "hat." A little later the pupil is given a pencil and is asked to make these same black marks which have the name "hat." The psychological series consists of, first, visual stimuli from the form of the letters to the visual centers in the brain, second, the establishment of connections in the brain between the visual centers and motor centers for the hand, and third, redirection of impulses from the motor centers to attempt to write the word "hat." Then from the muscular movements of the hand and arm, made more or less by trial and error, kinesthetic sensations are derived and associated in the mind with the visual stimuli of the word "hat." These two sets of sensations become associated and direct the motor responses in carrying out the writing act. Still later the pupil is given (1) the visual or auditory stimulus, "If you buy a pencil for three cents and give the clerk five cents, how many pennies should you receive back?" which (2) arouses a variety of association processes between various numbers such as multiplication, division, addition, or subtraction, and out of the mass of associations one is selected, namely, five minus three equals two, and this in turn (3) directs the impulse to the motor centers to say "two." All learning, even including reasoning, is probably of the same fundamental type. The only difference is that there are more elements involved in each of the parts of the three-series connections and that, owing to the larger number of elements aroused, a selecting or picking out of certain elements rather than of others takes place. Learning facts of history, economics, or science may be described in the same general schematic manner. The facts are either read in a book, heard spoken by the teacher, or observed directly. These sensory impressions are associated, dissociated, and combined in various ways, which in the course of time usually lead to some form of reaction either of speech or of larger muscular activity.

The preceding examples of school learning depend for the most part upon simple associations, that is, upon the law that things experienced together or in close succession tend to come back together. Thus after a certain number of repetitions the sight of certain black marks will set off promptly the reaction of speaking the word "hat." But in reality the process of learning is almost never as simple as this. While it is true that association bonds must be set up between situations and responses, a single situation is almost certain to present to the mind of a child of school age a multiplicity of aspects. As a consequence we find, instead of a single bond joining the response to the situation, a number of bonds each joining the response to a different part of the situation. Thus the word "triangle" may be associated with an equilateral triangle of red cardboard, $^1/_8$ inch thick, 8 inches on each side, showing a dull gray edge and weighing one ounce. Innumerably different combinations of bonds may accomplish this. There may be, for example, a major bond leading to a reaction to the redness, a secondary bond connecting the reaction with the size and symmetry of the sides, and minor bonds emanating from the thickness and color of the edge, while the weight and texture of the paper, the shade of red and, most important of all, the number of sides and the angles may not emerge from the complex at all. Clearly such a set of bonds would be worse than useless in the presence of a right-angled triangle indicated in a book by black lines on a white page with the angles labeled by letters and with a base ½ inch long. By dint of numerous experiences with a variety of triangles and with the help of the teacher who points out the essential three-straight-sidedness, the characteristics peculiar to a triangle finally emerge more or less clearly from the complex and become associated with the various reactions appropriate to "triangle." The false bonds are either destroyed or greatly weakened. Association is still the basis of the process, but there is in addition the dissociation from one another of the various characteristics which make up the complexes called objects. This is conveniently spoken of as learning by analysis and abstraction. When complete, the process of analysis and abstraction, which makes possible the reaction to parts of situations rather than wholes, clearly is an enormous advance over simple associative learning. One association thus properly attached to the significant part of a situation may function without any further effort in a great variety of similar but otherwise entirely novel situations. This is probably the essence of reasoning. But again we must note that the process is rarely so simple as has been outlined. Rarely are the preliminary analysis and abstraction so complete that a reaction is transferred without delay to a very novel situation. Besides, the attention may be distracted from the often subtle and inconspicuous but significant element in the new situation by the novel and striking but irrelevant features. Sometimes some of these irrelevant features touch off a reaction which is entirely inappropriate. For example, all but the very brightest pupils in a class, which has learned to compute with facility the area of triangles from printed problems and diagrams in a book and which knows how to measure accurately a straight line, would be completely at a loss to know what to do if given a 66-foot tape measure and confronted by an area of ground in the shape of a scalene triangle measuring 4 by 7 by 10 rods, covered with flower beds in a setting of greensward and surrounded by an ornamental iron fence three feet high. The sagacious few, if well taught, will react unerringly to the significant element in the situation. Such are the processes involved when a child has learned to isolate the significant element from the situation: "If you buy a pencil for three cents and give the clerk five cents, how many pennies should you receive back?" so that it touches off the subtraction reaction rather than that of addition, multiplication, or division.

Common and Special Elements in the Learning Process. If we grant that the stimulus-association-response series is the schematic pattern of learning, it will be convenient to discuss the psychology of learning in two parts: first, the psychology of learning in general in which the elements common to various types of learning will be examined; and second, the psychology of the learning of school subjects in particular in which the special elements and processes peculiar to each type of learning will be examined.

Practically all types of learning of whatever sort, have certain processes in common. They have in common certain elements of sensation and perception which are involved in the reception of the stimuli in any sort of learning. Associated bonds are formed in certain fundamentally similar ways no matter what the mental processes are between which the bonds are formed. Retention, assimilation, analysis, abstraction and generalization have certain uniform characteristics. Likewise the redirection and reaction processes follow certain general principles. But on the other hand, each type of learning has its own special sensory material presented and perceived in its own particular manner; it has its own special bonds which must be formed between its peculiar elements; and it has its own type of reaction occurring in its own individual way. Thus in learning to read, the series is, first, visual-auditory stimuli of words and groups of words; second, association of visual-auditory impressions and the memory of the objects which they represent; and third, the response of speaking. In learning to spell, the series is, first, visual-auditory stimuli of letters in a certain succession; second, association of the stimuli in their particular orders; and third, response in speaking or writing.

The Psychology of Learning in General versus the Psychology of Learning of School Subjects. The procedure of learning in general can probably not furnish the process and technique of the learning of school subjects. The process, the technique, and the economy of the learning of school subjects must be worked out experimentally in detail for each particular subject. All that the procedure, technique, and economy common to all types of learning can do is to furnish a broad general background. Thus the substitution experiment in the author's *Experiments* ('17), Chapter X, furnishes certain conclusions with regard to the distribution of time according to which the establishment of the bonds involved in this particular type of learning can be made most economically. This experiment shows that it is more profitable to work at the task 10 minutes twice a day than 20 minutes once a day, or than 40 minutes once every other day. It would, however, be folly to attempt to generalize from such an experiment merely and to say that it is better to teach writing 10 minutes at a time twice a day than 20 minutes once a day, or than 40 minutes every other day. All that the experiment indicates is the general principle that short periods of work distributed at certain intervals of time are productive of greater progress in learning material of this sort than longer periods distributed at longer intervals. What the length of the most economical periods and the intervening intervals would be in the case of writing, reading, or any other school subject, cannot be inferred on *a priori* grounds from a general principle, but must be determined in detail for each particular type of material and for the particular conditions under which the learning must take place. All that the general principle can do is to point the way to a more or less probable solution, but the particular direction and the course of the path must be determined from further observations. The factors and laws of the mind as set forth in general psychology can therefore not be carried over bodily into the psychology of a particular type of learning. General psychology can furnish its experimental technique and its fundamental laws which will serve as guides in the development of the psychology of special types of learning. From this point of view, the psychology of school subjects and the pedagogy of these subjects resulting therefrom, which is likely to be the only sound pedagogy worthy of the name, are as amenable to experimental attack according to rigorous, scientific procedure, as the problems in other fields of psychology have been amenable to the technique of experimental methods.

Program of Problems. According to our analysis, then, the following problems result:

A. The psychology of learning in general.
  1. How are the stimuli received?

a. How do sensory defects interfere?
b. What are the factors and conditions of observation and perception?

2. How are they associated, analyzed and combined in the mind?
   a. What is the rate and progress in the formation of associative bonds?
   b. What factors and conditions promote or hinder the most economical formation of the bonds?
   c. What are the effects of one set of associations upon other sets of connections? (Transference of training.)
3. How are they redirected into responses?

B. The psychology of school subjects in particular.
   1. What are the specific psychological processes involved in the learning of each particular subject?
   2. How may the capacities in each subject be measured?
   3. What factors and conditions promote or retard the learning in each particular subject?

CHAPTER IX

THE RECEPTION OF STIMULI: A. SENSORY DEFECTS

Importance of Normal Sense Organs. The first point of contact with external stimuli is through the sense organs. It is obvious that this point of contact should be as perfect as possible so that the stimuli which are to furnish the material for learning, may be received as accurately and completely as possible. The eye and the ear are the most important avenues of information. Defective eyes and ears necessarily make school work difficult and disagreeable. A child with defective eyes or ears misses a great many things which the normal child can see clearly or hear distinctly. The sad aspect of the matter is that a great many children have sensory defects of which neither they themselves nor their parents or teachers are aware. They assume that every one else sees or hears just as they do and consequently their attention is not called to it. These defects, however, often become serious and remain as permanent hindrances throughout life. Many of them, if discovered early, may either be kept from becoming aggravating or be reduced very considerably. Furthermore, sensory defects, and in particular visual defects, have far-reaching consequences upon a child's life as a whole through the production of headaches, nervousness, and dislike for school work. It is even claimed that truancy is in some instances indirectly traceable to visual defects on account of the dislike for school work produced by them. Dr. J. H. Cliborne has given a retrospect of his own school days during which he suffered from visual defects, as follows:

"I now know I have always carried about 1.50 diopters of hypermetropia; in my very early days, possibly more. Books and school were to me a nightmare, a source of unutterable disgust. I drove myself to my tasks with the scourge of duty; I never took one moment's joy or pleasure in the acquisition of knowledge, unless it was the satisfaction of a task accomplished or conquest gained. I have no memory of a sense of pleasure connected with my studies at school or college. The only pleasant memories I have are those connected with outdoor sports, or facts gained through observation, or in the lecture-room through my ears; and from my boyhood I could never understand why we were forced to read from books all that we learned.

"Early in life I pondered over the easiness of the task of those who never sat at the feet but who followed the tracks of the peripatetic philosophers. Verily, my school and college days would have been a joy to me had my ears and my distant vision been my means of acquiring knowledge; and yet I never had a headache in my life at school nor in after years until after the commencement of presbyopia. I was nervous to the point of madness at times, and the more nervous I was the more diligent I became, and the nearer I put my nose to my book. I have frequently observed that my right eye was crossed after prolonged study, or after a long written examination; this was also at times observed in my case by a fellow-student. That the difficulty lay in my hypermetropia I have no manner of doubt. I had inherited a love of learning, I felt sure, and I had a right to the assurance, and my hatred of close application was a mystery to me. I created a frown by my accommodative strain, which has ever been a part of me. Prolonged application to books would be followed often by sleeplessness or violence in the field at play. I learned for these reasons the art of complete concentration, but at what an expense of nervous energy." (Swift '12, pp. 94–95.)

As concrete illustration of the manner in which undiscovered visual defect may interfere with school work, we may cite the following case from the psycholgical laboratory of the University of Pennsylvania:

"On a certain afternoon in March, 1896, Miss Margaret T. McGuire, a grade-school teacher in the Philadelphia public schools, went to the psychological laboratory of the University of Pennsylvania, accompanied by a lad of fourteen—a well mannered, intelligent lad, industrious in his school work; one of the favorite pupils, in fact. Yet this lad was the 'bogy' of the teachers who for seven years had had him in their classes: he was a chronic bad speller. This does not mean that he misspelled some words sometimes. He misspelled every word always, and did it in the same careful and serious manner with which he recited the history lesson he loved. His reading was as bad as his spelling; he was absolutely incapable of getting through a single sentence correctly, *a*, *an*, *and*, *the*, and a few three-letter words being the net result of his seven years' schooling. He read *saw* for *was*, *water* for *weather;* wrote *htat* for *that*, *soas*, for *soap*, and other picturesque combinations of the sort in endless variety. His case seemed hopeless. . . . Dr. Witmer made a long examination, the result of which was the discovery that Charles Gilman had an ocular defect never, in all these years, so much as suspected by either his parents or his teachers: *at the distance of about three feet the boy* *saw everything double:* 'he lacked the power to direct the two eyes coordinately upon the same point in space, the left eye looking a little higher than the right.' A page of ordinary print was thus a blur; whenever he attempted to write, the words doubled under his pen. Curiously enough, he had never mentioned this peculiarity—he seemed to think it the natural process of vision. And he had repeated three whole years of school on this account alone. . . . He was fitted with glasses and later operated upon; then for the first time in his life the printed page and the words he was tracing with his pen were clear. But his reading and writing and spelling were just as bad as ever! The oculist had removed the *defect*—he had not removed the *effect* of the defect: that was in the boy's mind. And it was here that the psychologist came to the rescue by showing just what the effect was and how to remedy it.

"Now, it is an obvious truism of daily life that in order to recognize a thing when we see it again, we must have seen it, at least once, clearly and distinctly: a mental image of it must have been left in the mind. Reading is simply a rapid-fire recognition process by means of the stored mental images of words. Charles Gilman had no stores of images of words, for he had never seen any—he had seen only blurs of words. He was even worse off than the child just groping its way through the primer, for he had to unlearn the blurs he had patiently acquired through those seven years when nobody knew what his trouble was; then word by word, he had to restock his mind with the images of words shown him through his glasses. . . . In spite of this handicap, the boy learned to read, write, and spell, and was finally graduated from the grammar school only three years later than he should have been; which was better than not being graduated at all." (Carter '09.)

Types of Visual Defects. The most common forms of visual defects are myopia, hypermetropia, astigmatism, strabismus, and color-blindness.

Myopia, or near-sightedness, is usually due to the fact that the eyeball is too long and consequently the rays of light entering the eye are focused at a point somewhere in front of the retina. As a result the rays of light are again spread out when they reach the retina and therefore do not form a distinct image.

Hypermetropia, or far-sightedness, is usually due to the fact that the eyeball is too short and so the rays of light entering the eye are not sufficiently refracted in order to form a clear image when they impinge upon the retina. The image would be formed at a point back of the retina if the rays of light were extended. In some cases myopia and hypermetropia may be due to improper refraction of light by the crystalline lense or the cornea. These abnormal conditions cause a considerable strain upon the muscles of the eye which attempt to accommodate the lens in order to form as clear an image as possible.

Astigmatism is due to the fact that the curvature of the cornea or of the lens, usually of the former, is not the same in all meridians. The result is that if the lens is accommodated for rays of light from some parts of the field of vision, it will not be accommodated for rays of light from other parts of the field of vision, and consequently a portion of the field of vision will be distinct and another portion will be blurred. The curvature of the cornea or of the lens in the case of astigmatism may be compared to that of an eggshell. It is different in different directions whereas in the normal eye it is similar in all directions like that of a sphere.

Strabismus refers to the lack of perfect balance in the external muscles of the eyes, so that the two eyes do not focus upon the same point at the same time. This difference in point of focus may, of course, vary anywhere from perfect coincidence, as it should be in the normal eye, to a very large deviation, commonly known as cross-eyedness, which can be observed readily by looking at a person's eyes. The history of Charles Gilman cited on preceding pages was a case of strabismus.

Color-blindness consists in the confusion of certain colors, nearly always red and green. Confusion of the other two fundamental colors, yellow and blue, almost never occurs. The cause of it is more or less speculative and may be due to the absence, or the improper functioning, of the color elements in the retina. Color-blindness is obviously a drawback in any type of school work in which colors are concerned, such as drawing, map work, domestic science and manual arts and all scientific studies in which color discrimination is involved.

Causes of Visual Defects. Visual defects are due in general to two causes, (1) heredity, and (2) the strained use of the eyes for fine distinctions at close range, particularly under poor illumination. Color-blindness and strabismus are probably inherited in nearly every instance. The other types of defects which relate to the formation of the image are proably due in part to hereditary conditions in the sense organs and in part to overstrained use of the eyes. Reading, which has been the great promoter of civilization, has also been, in a certain sense, a deteriorator of the eyes through the strain put upon them by the fine distinctions that must be made at close range and at a tremendously rapid rate. Durr [as reported by Whipple ('10, p. 139,)] has attempted to explain the great prevalence of myopia in Germany as compared with other countries by the excessive demands made by the German school system. He has estimated the number of hours devoted to study and to exercise by the typical boy during the years 10 to 19 in different countries as follows:

	Hours study	Hours exercise
Germany	20,000	650
France	19,000	1,300
England	16,500	4,500

Most hygienists maintain that myopia is an acquired condition whereas anatomists are more inclined to regard it as an inherited condition. Cohn (Whipple '10, p. 139) reports that myopia occurred in gymnasia in the following increasing percentages during the six years of study: 12.5, 18.2, 23.7, 31.0. 41.3, 55.8.

Frequency of Visual Defects. Recent years have brought a considerable number of investigations as to the percentage of children with defective vision. Whipple made an examination of the vision of 1,000 white and 100 colored children in Jefferson City, Missouri, in which he found the following percentages of visual defects:

TABLE 3[illegible]. After Whipple

Visual defects among 1 000 white and 100 colored children in Jefferson City, Missouri

	White	Colored
Defective vision (Snellen test)	36.5%	19%
" " (one eye)	13.8	17
" " (both eyes)	22.7	12
" " (first 3 grades, 147 pupils)	29.4	
" " (high school, 116 pupils)	40.5	
Pain after using eyes in study	29.5	34 [1]
Probably needing glasses	41.0	
Wearing glasses when examined	3.8	5
Cross-eyed	3.0	2

Taussig ('09) has summarized the percentages of visual defects in various cities in different countries as follows:

[1] This percentage is probably too high since it was discovered that the colored children took peculiar pride in reporting headaches because they seemed to consider it a sign of intellectual keenness.

TABLE 31. After Taussig ('09)

Heidelberg, Germany (1870)	35.0%
Edinburgh, Scotland (1904)	43 2
Dunfermline, Scotland (1907)	17 0
Cleveland, well-to-do district (1907)	32 4
" congested district (1907)	71.7
Massachusetts, except Boston (1907)	19 9
Boston and environment (1907)	30 7
Boston (1908)	23 0
New York City (1906)	31 3
New York City, Borough of Manhattan	10 2
Chicago (1909)	19.4
Jefferson City, Mo., either eye (1908)	36 5
" " " both eyes (1908)	22 7
St. Louis County, Mo., either eye less than 20/20	30 6
" " " " both eyes less than 20/30 (1909)	14.3
" " " " both eyes less than 20/40 (1909)	2 8

Additional results from other cities as reported by Gulick and Ayres ('08, p. 83) are as follows:

TABLE 32

Bayonne, N. J	7 7%
Camden, N. J. (1906)	27 7
Milwaukee (1907)	14.7
Minneapolis (1908)	23 9
Pawtucket, R. I. (1901)	11.1
Utica, N. Y. (1897)	10.9
Worcester, Mass	19.1

The large variations in the percentages quoted for different cities probably do not represent actual differences in the prevalence of visual defects. They are probably due mainly to differences in standards adopted by various examiners. Whether an eye is reported as defective or not is, in the milder forms of defect, an arbitrary matter. From the purely mechanical standpoint, an absolutely perfect eye is undoubtedly very rare. Therefore the matter resolves itself largely into the question as to whether the deviation from perfection is sufficient to interfere appreciably with normal distinct vision. As a general statement we may say, according to the quoted tables, that approximately 25% to 33% of the school children have visual defects sufficiently serious to demand some attention. Color-blindness fortunately is relatively rare. It is estimated that 4% or 5% of men and less than 1% of women are color-blind.

Visual defects seem to increase measurably with successive years in school. Thus in Whipple's table the percentage of defect among high school pupils is 11% higher than that among pupils in the first three grades. An extensive comparison was made by Gulick and Ayres ('08) in New York City which showed the following percentages grade by grade:

TABLE 33

Grade	Per Cent of Visual Defect
2	20.2
3	21.9
4	25.8
5	24.8
6	24.5
7	26 9
8	32 3

Remedial Measures for Avoiding Visual Defects. The first and probably most important suggestion is the examination of the eyes of pupils at least once in two years or preferably once a year. This would serve as a means of discovering the visual defects so that measures could then be adopted for the proper care of those pupils suffering from them. Excellent results have been shown in various cities in which general sensory examinations have been introduced. Taussig reports that in Boston the percentage of visual defects dropped from 30.7% to 23% and in New York City from 31.3% to 10.2% as a result of the introduction of visual examinations. The schools require compulsory attendance, but they have not taken sufficient steps to make it possible for the children to remain in school to their greatest profit.

In the next place, the proper lighting of schoolrooms is highly important. This matter is being taken care of, however, at the present time by school architects in a much more thoroughgoing manner than was formerly the case. Many of the older buildings are poorly arranged and wretchedly lighted. The amount of window area to floor area usually recommended as satisfactory, is approximately one to five or six. As typical of the inadequacy of the lighting in older school buildings we may note that in a test of the amount of light at different desks in one of the old school buildings in Madison, made by Cohn's light tester, showed that in a room of five rows of seats the two rows next to the inside wall showed a decided insufficiency in illumination.

Numerous other precautions may readily be exercised by the school to avoid the aggravation of existing visual defects. Such measures are the breaking up of the school program to spread the severe use of the eyes, reduction of close work in the early years of a child's life, prohibition of work in poor artificial light, adjustment of the size of the desk to the size of the pupil, the instilling of the habit of resting the eyes even for the short interval of a few minutes in the midst of eye-straining work. The proper printing of books and the use of appropriate paper is being looked after by publishers much more carefully to-day than formerly.

Frequency of Auditory Defects. The presence of a greater or less amount of deafness in one or both ears often interferes very considerably with normal school work. Whipple has reported for the 1,000 white and 100 colored children in Jefferson City, Missouri, auditory defects as follows:

	White	Colored
Defective hearing (whisper test)	7.7%	7.0%
" " (one ear)	6 4	4.0
" " (both ears)	11.3	17.

Taussig ('09) has reported the following percentages of defective hearing in various cities:

TABLE 34

Edinburgh, Scotland (1904)	12.2%
Dunfermline, Scotland (1907)	4.0
Cleveland, well-to-do district (1907)	5.2
" congested district (1907)	1 8
Massachusetts, except Boston (1907)	5 8
Boston and environment (1907)	7 7
Boston (1908)	7 6
New York City (1906)	2 0
New York City, Borough of Manhattan	1 0
Chicago (1909)	2 7
St. Louis County, Mo., either ear defective (1909)	7 3
" " " " both ears seriously defective (1909)	2.2

Additional results reported by Gulick and Ayres ('08) are as follows:

TABLE 35

Bayonne, N. J.	2.5%
Camden, N. J. (1906)	4.1
Minneapolis (1908)	7.7
Pawtucket (1901)	4.3
Utica, N. Y. (1897)	6.6
Worcester, Mass	6.6

As a general statement we may say that approximately 5% to 10% of pupils suffer from defective hearing of a sufficiently serious character to interfere with their school work and to require medical attention.

Effects of Sensory Defects upon School Work. A number of inquiries have been made to determine the amount of hindrance which sensory defects have upon the proper performance of school work. Dr. Cornell has reported the following results for three schools in Philadelphia showing the average school marks of normal and defective pupils:

TABLE 36. After Cornell

Allison School—219 children, both sexes, 6 to 12 years old

	Average
Normal child	75
Average child	74
General defectives	72.6
Adenoids and enlarged tonsils	72
Deaf	67.2

Ninth Street Primary School—84 children, both sexes, 6–10 yrs. old.

	Language	Arithmetic	Spelling	Average
64 cases normal children	72.9	75.5	75.4	74.6
84 cases average children	70.5	74	72.8	72.4
21 cases general defectives	63.3	70	64.8	66
8 cases adenoids	60.0	66.7	65	63.9
No cases deaf.				

Claghorn School—252 children, both sexes, 12 to 15 years old.

	Language	Arithmetic	Geography and History	Average
179 cases normal children	74.4	72	76.6	74.3
252 cases average children	72.7	70	76.5	73.1
73 cases general defectives	71.4	65.1	76.2	70.8

Whipple in his study in Jefferson City found that among pupils of good vision, 26% did unsatisfactory work and among pupils with defective vision 38% did unsatisfactory work. Smedley reported that in Chicago there were 18% with auditory defects among pupils above grade, and 25% among pupils below grade. It is difficult to say how much of the scholastic inferiority of those having sensory defects is due to these defects and how much is due to inferior native ability. It has been shown by Pintner and Paterson ('16) that deaf children are fully three years behind normal children in learning the digit-symbol test which, as given by them, does not depend upon the use of language. Goddard ('14) in an intensive investigation found blindness perceptibly associated with feeble-mindedness, and deafness with neuropathic taint.

Musical Discrimination. Besides the various degrees of deafness, the sense of hearing is of direct interest to the school from the standpoint of musical instruction. The ability to sing, and to some extent, to appreciate music, depends in part upon the accuracy of the discrimination of pitch. This ability varies widely among people, and considerable inaccuracy in musical discrimination is a distinct difficulty in learning music. The results of recent experiments seem to indicate rather definitely that accuracy in pitch discrimination can probably not be improved by practice or by teaching. The voice in singing and the hand in playing the violin, are guided by the accuracy of the ear. If the ear is not accurate the individual is unable to guide his voice or his fingers with a precision necessary for the production of music.

Seashore ('01), who has studied this problem carefully, has suggested that pupils whose discrimination is two vibrations or less have a sufficiently fine ear to become musicians; pupils whose discrimination lies between 3 and 8 vibrations, which includes the large majority of people, have a sufficiently accurate discrimination for ordinary musical instruction and enjoyment; pupils whose discrimination lies between 9 and 17 vibrations should have musical instruction only if they have special inclination or desire for it and singing in school should be optional for them; and finally, pupils whose discrimination lies at 18 or above, should probably not be required to study music or to attempt to produce music.

Physical Defects. Beside the sensory defects there are several types of physical defects whose frequency is high and whose interference with school work is serious. The prevalence of these defects is indicated in the following table, adapted from Gulick and Ayres ('13, p. 38):

TABLE 37

	Percentages of children having defects of				
	Teeth	Throat	Nose	Glands	Others
Boston (1912)	0 8	22 6	8 2	12 4	17.7
Chicago (1910)	36 5	20 2	8.1	13 8	8 9
Cleveland (1910–11)	32 4	15 3	11 6	9.8	7.5
Newark, N. J. (1910–11)	29 3	18 9	12 3	17.1	13.7
New York (1911)	59 0	15 0	11 9	.	4 6
Oakland, Cal. (1910–11)	48 1	35 8	18 1	8 9	3 0
Pasadena, Cal. (1909–10)	30 5	5 0	5 0		6.0
Rochester, N. Y. (1910)	44 8	29 4	17 4	8 5	17 0
St. Louis (1910-11)	52 0	17 7	4 8	.	1 2
Average	41.5	20.1	10.8	11.8	8 8

CHAPTER X

THE RECEPTION OF STIMULI: B. PERCEPTION AND OBSERVATION OF THE SENSORY MATERIAL

Problems. Next to the normal operation of the sense organs come the actual perception and observation of stimuli as presented to the sense organs. What material you learn and how you learn it depends upon what and how you perceive or observe. Obviously what you perceive or observe depends upon the normal operation of the sense organs. That has long been recognized. But it has not been so fully recognized that it depends also upon the mental apprehension of the sensory stimuli. Perception and observation do not depend alone on what is presented to the sense organ, but also upon how the stimuli are taken into the mind. The specific problems involved in it are:

(1) How accurate is the observation of stimuli?
(2) How large is the range of stimuli observed at a given time?
(3) How may the accuracy and range of observation be improved?
(4) How are the stimuli interpreted?

Accuracy of Observation. Much of the difficulty in learning a given material is due to inaccuracy and error in the observation of the material. The word to be spelled, the letter to be written, the plant to be described, the experiment to be reported, the map to be drawn, may all be done and learned incorrectly in part because they are perceived and observed inaccurately and incompletely.

Recent experiments on the "fidelity of report" have called attention to the prevalence and nature of inaccuracies in observation. The observations are not made with sufficient care and attention to impress a faithful image of the object upon the mind. Inaccuracy in perception and infidelity in report have hardly been known to exist and therefore have not been sufficiently guarded against because, under ordinary conditions of learning in school as well as under ordinary conditions in life, there is seldom an opportunity for comparing directly the observations as received in the mind with the original stimuli as actually presented to the sense organs.

The recent investigations on the fidelity of report have been carried out by presenting to a group of observers a series of events either in the form of real actions or more commonly by means of a picture. The picture is exposed to the observers for a short period, say thirty seconds or a minute, after which they are asked to write a report of their observations. This is usually supplemented by an interrogatory report consisting of answers to questions regarding the picture. Experimental inquiry into these matters has been stimulated primarily from the practical importance of determining the reliability of witnesses in court. Incidentally, the results have an exceedingly significant bearing upon the accuracy of observation involved in learning. The main results of these experiments have been summarized by Whipple ('10, pp. 304 ff.) as follows:

"The chief single result of the Aussage psychology is that an errorless report is not the rule, but the exception, even when the report is made by a competent S (subject) under favorable conditions. Thus, in 240 reports, Miss Borst found only 2% errorless narratives and 0.5% errorless depositions. The average subject, when no suggestive questions are employed, exhibits a coefficient of accuracy of approximately 75%.

"Generally speaking, attestation does not guarantee accuracy: on the contrary, though the numbers of errors is nearly twice as great in unsworn as in sworn testimony (according to Stern, 1.82 times, according to Borst, 1.89 times as great) there still remains as high as 10% error in sworn testimony.

"Reports of children are in every way inferior to those of adults: the range is small, the inaccuracy large, and, since the assurance is high, the warranted assurance and reliability of assurance are both very low. During the ages 7 to 18 years, the range, especially the range of knowledge, increases as much as 50%, but the accuracy, save in the deposition, does not increase as rapidly (20%). This development of capacity to report is not continuous, but is characterized by rapid modification at the age of puberty.

"The one factor that more than others is responsible for the poor reports of children is their excessive suggestibility, especially in the years before puberty.

"Not all the features of the original experience are reported with the same frequency or with the same accuracy: there is, rather, a process of selection, both in the process of observation, and also, probably, in memory and in the formulation of the report. In general, we may say that persons and their acts, objects, things, and spatial relations are reported with considerable accuracy (85–90%), whereas secondary features, especially quantities and colors, are reported with considerable inaccuracy (reports on color have an error of from 40–50%).

"All authorities agree that the use of interrogatory, whether of the complete or incomplete form, increases the range and decreases the accuracy of the report.

"The introduction of leading or suggestive questions very noticeably decreases the accuracy of report for children, and, unless the conditions of report are quite favorable, even for adults. The greater suggestibility of children is shown by Stern's results in which the inaccuracy of boys and girls, ages 7 to 14 years, was from 32 to 39%, as against 10% inaccuracy for young men aged 16 to 19 years."

The reasons for inaccuracies in reports of observation must be sought in various directions. The chief ones are (1) insufficient attention to the material observed so that only a vague impression is produced which may easily be disturbed and modified by other stimuli, (2) meagerness of ideas and experiences with which the observed material may be connected and interpreted, (3) low retentiveness of the impressions so that other impressions can readily distort them, (4) faint imagery in terms of which to picture the objects, (5) lack of conscientiousness in keeping apart the observed and the inferred items as a result of which missing or doubtful elements are supplied by unconscious inference, (6) the effect of suggestion through which the mind is prone to seize upon any slight hint and to fit it into the story.

The Range of Observation. The amount of material observed by different persons within a given limit of time varies over an astonishingly wide range. Individual differences in the capacity for apprehending stimuli from the outside world are probably as large as those in other mental abilities. One person may report several times as much of a scene or series of events as another. The pupil with a wide scope of apprehension and observation has a tremendous advantage over one with a narrow scope of observation. This range probably depends fundamentally upon the span of attention, quickness of assimilation of items, retentiveness, and previous knowledge about the facts to be observed. The span of attention as measured by rapid exposure methods, ranges in normal persons from three or four objects to eight or nine. Such experiments are carried out by exposing momentarily before the observer cards with varying numbers of objects and by recording the number of objects noticed.

Some years ago a number of interesting investigations were made by G. Stanley Hall ('83) and others to determine the range of observation and knowledge of pupils entering school. The returns showed a surprisingly narrow range of information and the great extent to which it was determined by the environment in which the child lived. The educational value of such inquiries has been to emphasize the importance of a wide variety of immediate contact and experience with real objects by observing, manipulating and using them.

As a concrete illustration, the author made the following simple experiment to indicate the range and fidelity of such observations as would be made in a class in biology. Some plants in a jardinière were exhibited before a class of thirty-nine students for thirty seconds with instructions that they were to observe them as carefully as possible and that they would be asked immediately afterwards to record their observations. The shortest report (A) was only fifty-seven words in length, and one of the longest and best (B) was 131 words in length. One of the most erroneous ones is given under (C). B is more than twice as complete as A, and fully as free from error. C is quite typical with regard to kind and frequency of misstatements. The erroneous portions are italicized and the corrections are given in parentheses. All three persons had had an elementary course in botany.

Report A:

"There was a high upright geranium plant having no buds but broad leaves. This was surrounded by low plants with drooping stems and bearing many pointed, small leaves. The leaves had a pinkish center, surrounded by a pale green *band*, the contour very irregular and the general effect bushy. There were no buds on these plants either."

Report B:

"From a brown jardinière arose one stalk of a geranium bearing nine big green leaves. The leaves spread out in all directions and are round in shape with large scallops. Lower was another plant with much smaller leaves and more bushy. It had three large branches, one leaning over the pot on each side and one across the front. The leaves of this plant were more oblong in shape, rounding at the base and reaching a point. The color was a pale pink in the center and back to the base, shading to a *deep red or wine* (pink) color towards the tip end and the whole edge of the leaf was green. The stem was much more delicate than the stem of the geranium and the leaves were much more numerous."

Report C:

"The plant was in a brown bowl. There were varieties of plants. The one had one large stalk with *five* (nine) branches growing out of it. The leaves were large and *heart-shaped* (rounding). The other plant was lower and drooped, and had more leaves. The leaves were *oblong* (heart-shaped) and smaller than the leaves of the first plant. They had a center of *very light green* (pink) and were outlined by darker green. The leaves were *smooth* and *glossy* (velvety). The leaves of the tall plant were notched."

It is obvious that the reporter of B has a great advantage over either A or C both in the quantity and in the quality of his observations. In a given period of time B will learn much more material and assimilate it in more correct form than either A or C. The objection might be raised that a report formulated after the observations have been made is bound to be erroneous and that it would be fairer to have all three persons stand before the plant and record their observations at the time. To this, however, we may reply that the range of facts observed and recorded in a given period of time would be just as wide and that observed items are not really assimilated mentally until they can be adequately thought out and expressed. The simple experiment here reported probably represents quite fairly the sort of things that occur in ordinary observation of material in learning.

Improvement in Observation. Granting the importance of a wide range and a high degree of accuracy in the perception of the material to be learned, what may we do to increase the accuracy, scope, and fidelity of observation? Probably the only advice to give at the present time is the rather obvious suggestion: Insist on greater accuracy and attention in observation. This may be accomplished by definite discovery of errors and inaccuracies in the observations themselves. Experiments have shown that many people, especially children, do not realize their inaccuracies and that calling attention to the discrepancies between objects and mental impressions of them results in a material reduction in unreliability. Whipple ('10, pp. 309 ff.) has summarized the experimental results on the possibility of improvement in observation by repeated tests with the same groups of persons as follows:

"Simple practice in reporting even without special training or conscious effort to improve, facilitates and betters the report, as is shown in Table 47, from Miss Borst. It will be noted that the tendency to oath and warranted tendency to oath are both particularly improved by practice, and that there is also an appreciable improvement in range, accuracy, warranted assurance, and reliability of assurance, whereas assurance and accuracy of assurance are scarcely affected. Similar practice-effects may be discerned in their deposition. From these results, it is clear that the several coefficients of report may vary more or less independently."

EFFECT OF PRACTICE UPON COEFFICIENTS OF REPORT (NARRATIVE) (BORST)

Number of Report (Test)	1	2	3	4	5
Range	39.0	39 0	42.3	40.3	42.0
Accuracy	86.6	87.7	92.9	88.2	90.0
Assurance	96.6	96.4	97.8	97.9	98.6
Warranted assurance	84.0	87.0	91.0	88.0	89.0
Reliability of assurance	87.5	89.4	92.6	89.8	90.3
Accuracy of assurance	97.0	98.0	98.4	98,6	99.2
Tendency to oath	43.0	59.8	62.8	61.9	72.1
Warranted tendency to oath	40 2	53.2	58.5	57.5	66.5
Unwarranted tendency to oath	2.8	6.6	4.3	4.4	5.6
Reliability of oath	93.0	88.8	92.5	93.0	91.7

NOTE: The effect of practice in these tests is somewhat obscured by the fact that the first and third tests were made after a 3-day, the others after a 9-day interval.

"The capacity of children to observe and report in a detailed and accurate manner may be improved by systematic training. This education may be best secured by appeal to zeal, interest, enthusiasm, or desire for improvement on the part of the child; more formal training of an intellectual type, e. g., suggestions for systematic observation, specific training in sense-perception, instruction designed to augment appropriate apperceptive-masses, etc., is much less effective.

"The inadequacy of the child's report is due, not so much to poor memory, as to the fact that he fails to perceive many features in the original experience, that he fails to put into words even what he does perceive, and especially to the fact that he is absurdly uncritical (his assurance, indeed, commonly reaches 100%)."

Tests such as these, but made with the material of school studies, would probably be very useful in bringing about more concentrated attention upon, and greater reliability in, observation. Thus a plant or a flower in a course in biology, might be exposed for a definite period of time to a class of pupils who would then be asked to write as accurate a description of the object as possible. This description could then be definitely compared, point by point, with the original object and in this manner the errors and inaccuracies would be discovered and noted. Difficulty in acquiring the material in various school subjects is no doubt traceable in larger part than we realize to the incompleteness and the unreliability of the perception of the material or stimuli to be acquired. It would be an experiment worth making to determine to what extent the difficulty of a pupil in learning to spell is due to actual incompleteness of the observation and perception of the letters in the words.

Interpretation of Stimuli. To a large extent, observation is interpretation. The same identical sense impressions are interpreted very differently by different observers. This may be demonstrated perhaps in extreme form in such tests as the one with the ink blots outlined in the author's *Experiments*, Chapter XIII. The first ink blot in that series signified to eight persons the following different things: map, bear, trees, lake, cloud, child, bat, man running. The same mental processes occur in a less variable manner in all kinds of observation. Incoming stimuli are interpreted by the association processes aroused in the mind. On this basis the traditional doctrine of apperception has been formulated and from it have been derived such pedagogical corollaries as, "Link the new to the old," or "Proceed from the known to the unknown." The theory of apperception has been very clearly expressed by James in the following statement:

"The gist of the matter is this: Every impression that comes in from without, be it a sentence which we hear, an object of vision, or an effluvium which assails our noses, no sooner enters our consciousness than it is drafted off in some determinate direction or other, making connections with the other materials already there, and finally producing what we call our reaction. The particular connections it strikes into are determined by our past experiences and the 'associations' of the present sort of impression with them. If, for instance, you hear me call out A, B, C, it is ten to one that you will react on the impression by inwardly or outwardly articulating D, E, F. The impression arouses its old associates: they go out to meet it; it is received by them, recognized by the mind as 'the beginning of the alphabet.' It is the fate of every impression thus to fall into a mind preoccupied with memories, ideas, and interests, and by these it is taken in. Educated as we already are, we never get an experience that remains for us completely nondescript: it always reminds of something in quality, or of some context that might have surrounded it before and which it now in some ways suggests. This mental escort which the mind supplies is drawn, of course, from the mind's ready-made stock. We *conceive* the impression in some definite way. We dispose of it according to our acquired possibilities, be they few or many, in the way of 'ideas.' This way of taking in the objects is the process of apperception. The conceptions which meet and assimilate it are called by Herbart the 'apperceiving mass.' The apperceived impression is engulfed in this, and the result is a new field of consciousness, of which one part (and often a very small part) comes from the outer world, and another part (sometimes by far the largest) comes from the previous contents of the mind." ('99, p. 157.)

The doctrine of apperception as here stated by James is simply a statement of the psychology of perception as usually accepted. The importance of the applications of the doctrine to teaching has perhaps been overemphasized in the educational writings of the recent past and in the pedagogical methodology that has been worked out in accordance with its corollaries. The applications of the theory as represented by the injunction, "Link the new to the old," is no doubt sound from the psychological side and useful from the pedagogical side as a general guiding principle. Illustrations of the principle would be the teaching of a topic in geography by connecting it up with the known facts of geography in the immediate environment of the child, or of teaching laws of chemistry by relating them to familiar facts and problems that have arisen within the child's own experience, or of teaching the spelling of a new word by pointing out its similarities to words already known, or of teaching forms of a foreign language by referring them to related forms in the language previously acquired. Such a procedure is unquestionably valuable whenever it can be employed. However, some of the enthusiastic advocates of the doctrine have been somewhat blinded to the limitations of it. If we regard learning as a process of establishing connections between elements of the learning-material, we can conceive of three possible kinds of bonds to be formed: (1) between two known elements which had previously not been connected, (2) between a known and an unknown element, (3) between two unknown elements. It is obvious that a great deal of learning consists in the formation of the third type of connections. The doctrine of apperception can apply only to the first and second type of connections and these can very probably be formed more readily according to the natural workings of apperception because some of the elements had previously been acquired.

Much of the discussion in favor of the doctrine of apperception is really based upon the greater practical value of the known and nearer at home, and upon the urgent need of knowing something about the immediate environment rather than about a distant time or territory, whose history or geography may be of little value to the child, than upon greater ease in the formation of bonds between a known and an unknown element. The main thing in education is not to proceed from the known to the unknown, but rather to acquire the unknown. If this can be done by linking the new to the old, well and good, but the chief object is the linking of the new. Much of the so-called proceeding from the known to the unknown or of the linking of the new to the old, is more or less fruitless, since it neither proceeds to the unknown nor links anything new. It usually consists of a reawakening of the known and of the old. Learning is fundamentally the acquisition of new sets of stimuli-association-response series. When a pupil first attempts to write, he must acquire new neural connections in securing control of his hands. When he begins to learn the meaning of printed symbols, he is confronted with new stimuli among which new connections must be established. Much of the so-called teaching according to the theory of apperception consists in setting up problems concerning things with which the child is already familiar and thus in arousing in him a desire to learn something new. This is, no doubt, good teaching, but the important part in learning is the new element to be sought and the new associations to be built up.

CHAPTER XI

THE RATE AND PROGRESS OF LEARNING

Problems. The chief problems to be considered are as follows:

(1) How rapidly are new associative bonds formed?

(2) Does the rate of acquiring new connections and new materials continue uniformly per unit of time or per repetition?

(3) Do variations in rate occur in a uniform manner?

(4) What causes will bring about variations in rate?

(5) Does the rapidity of learning occur in a similar manner in all types of learning?

Such problems as these may profitably be raised with regard to any sort of learning. If we consider the learning of a language we may ask, How rapid is the progress in acquiring the meanings of the words, knowledge and use of grammatical forms and ability to translate? Is progress uniform or are there times of rapid advance occurring in alternation with periods of little or no gain? What conditions will promote the learning of the language most effectively? If we could answer these and similar questions concerning any type of learning we would be able to control its progress far more economically than we are able to at the present time.

The Curve of Learning. The rate and progress of learning may be expressed in terms of the amount done per unit of time, or in terms of time required per unit of work. The relation between these two variables is represented by the curve of learning in which one function, usually time, is represented along the base line, and the other, usually amount accomplished, is represented along the vertical axis. Figure 36 represents a typical curve of learning in which progress is measured by the amount achieved per five minutes of time. It represents the rate of forming associations between numbers and letters in transcribing letters into numbers as specified in the author's *Experiments* ('17), Chapter X.

Characteristics of Learning Curves. Most of the experimental work on the course of learning curves has been done chiefly with various kinds of skills such as telegraphy, typewriting, tossing balls, mirror tracing, substitutions, and the like. Little has been done on the progress of analytical types of learning, on the advance in the acquisition of facts of a science or of the history of a country,

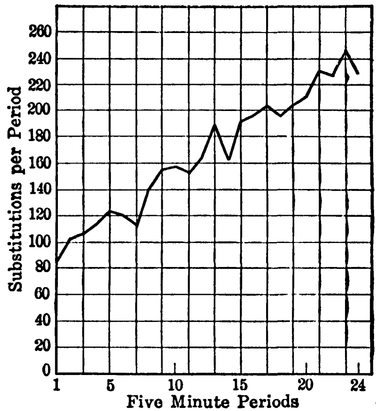

FIG. 36.—A curve of learning showing the progress of one person in learning to substitute numbers for letters in the experiments outlined in Chapter X, *Experiments in Educational Psychology.*

or on the rate at which a child learns to read or to write. Consequently most of our generalizations up to the present time have been based upon curves that represent the acquisition of skill.

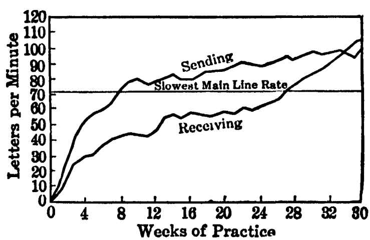

FIG. 37.—Improvement in telegraphy. Individual E. L. B. After Bryan and Harter ('97, p. 49).

Such curves seem to have in common two general characteristics, although it is doubtful whether they are universal in all types of learning: (1) An initial period of rapid progress, and (2) successive periods of no progress, or plateaus, followed by periods of rapid progress. Theoretically, a curve of learning may have two initial directions: (1) rapid progress followed by slower progress, or (2) slow progress followed by more rapid progress, that is, a convex or a concave form with all possible shapes between these extremes. The large majority of curves of learning derived to date are of the former sort. All the illustrations reproduced in Figures 37 to 41 have this general shape. In the substitution test referred to in Figure 36 the author found that among twenty curves obtained from as many individuals, thirteen were of convex form, six were practically straight lines rising from left to right, and one was of concave form. Hence, initial rapid gain seems to be a very common feature

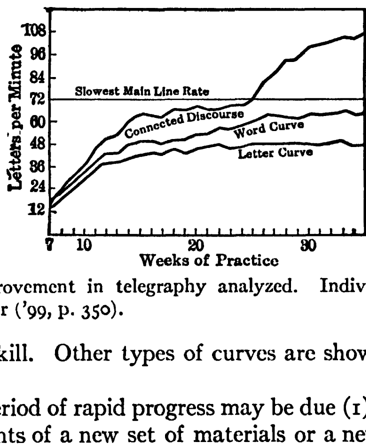

FIG. 38.—Improvement in telegraphy analyzed. Individual J. S. After Bryan and Harter ('99, p. 350).

in curves of skill. Other types of curves are shown in Figures 42 to 44.

The early period of rapid progress may be due (1) to the fact that the first elements of a new set of materials or a new set of associations may be picked up rather easily and quickly because of their simplicity, (2) to the probability that the first stage of practice in a new type of learning makes available various elements or activities already in the possession of the learner, (3) to the initial zeal in beginning a new task, (4) to the large opportunity for progress in the beginning, (5) to the physiological limit in many types of skill such as typewriting, mirror tracing and writing numbers for letters, and (6) to an absolute limit of the number of bonds that the task presents to the learner. Thus typewriting has a *physiological* limit in the rapidity with which the fingers can be moved in striking the keys. Progress cannot go on indefinitely at the original rate. Typewriting also has an *absolute* limit in the number of strokes to be learned.

Book explains the initial period of rapid progress thus:

"After what has been said our explanation of the general features of our curves can be brief. The first rapid and continuous rise is due to the fact that the learner is making progress along many different lines at once. Rapid strides of improvement are possible and made simulta-

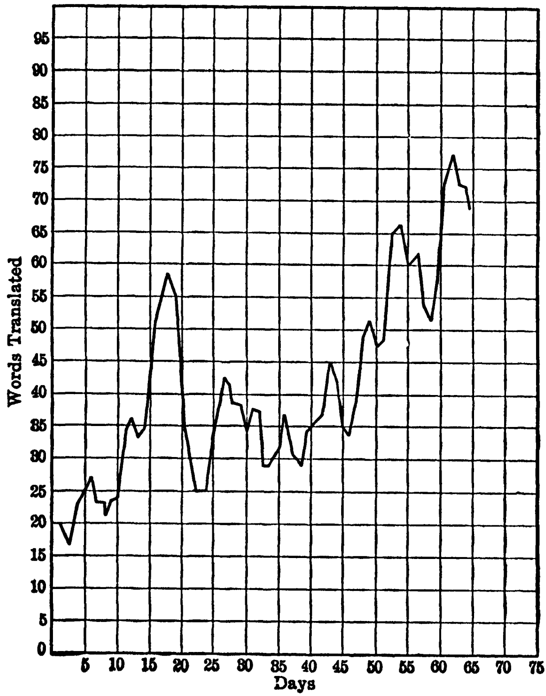

FIG. 39.—Progress in learning Russian. After Swift ('08, p. 198).

neously in every department of the work. The learner is not only forming and perfecting letter associations but syllable, word and phrase associations as well. He is simultaneously improving his method of dealing with every problem that the writing presents; locating the keys, directing and controlling his fingers, 'spelling' or initiating the movements, getting his copy, learning to deal with special difficulties, learning to keep attention more closely and economically applied to the work, etc. The curves will rise rapidly and continuously so long as many of these possibilities of improvement exist. As they grow less numerous the rate of gain will likewise decline until, as still more skill is acquired, a state is reached where most adaptations or short cuts in method have been made; fewer

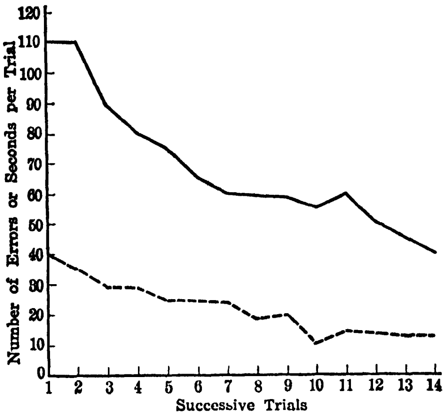

FIG. 40.—Improvement in tracing a star outline when seen in a mirror. Continuous lines represent reduction in seconds in successive trials. Dotted line represents reduction in errors.

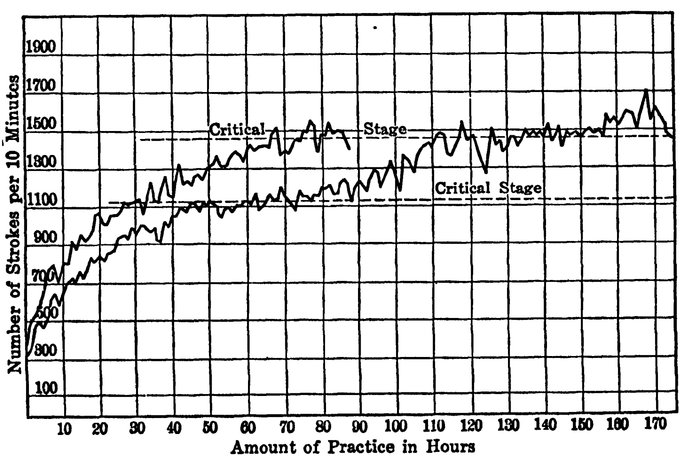

FIG. 41.—Improvement in typewriting by the sight method. After Book ('08 plate, opposite p. 22,

special habits remain to be developed; fewer adaptations are possible. Those possible have become harder and harder to make, because they must be made in the realm of higher habits where the learner has had less experience. Every man has had experience with the first stages of learning, but little with the later stages because most people touch lightly many things and are masters of nothing. There being now fewer adaptations to make, and the process of finally perfecting all the special associations being so gradual and slow, the learning curve becomes, as

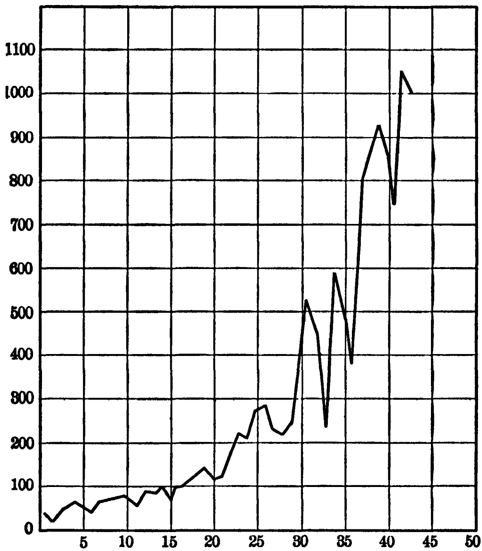

FIG. 42.—Progress in ball-tossing. The horizontal axis represents days. The vertical axis represents the number of balls caught. After Swift ('08, p. 174).

the expert stage is approached, almost horizontal. In the later stages of learning the sole gain must come from an occasional adaptation and from a further perfection of the present habits and methods of work." (Book, '08, pp. 99 f.)

Swift and Batson have each published curves based on the increase in skill in ball-tossing which purport to be of the concave type. A careful examination of the original data shows that the apparently concave form is in reality due in each case to peculiarities in the method of plotting. Fortunately some of the data published by Batson permit of being plotted strictly according to the principle laid down at the beginning of this chapter, namely, that vertical distances represent amounts of performance and horizontal

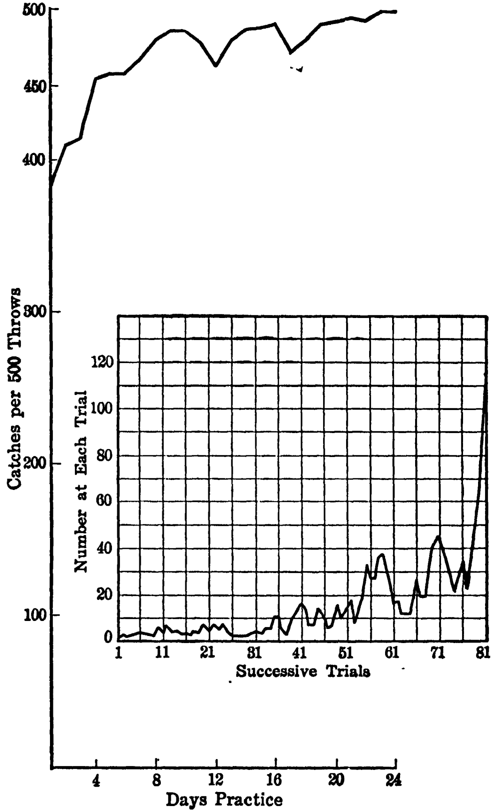

FIG. 43.—Lower graph shows Batson's original curve. Upper graph shows the reconstruction of his curve as stated in the text.

distances represent equal amounts of time or practice. This method of plotting yields the perfectly typical convex learning curve shown in Figure 43.

The types of learning so far investigated have been for the most part of a relatively simple sort. Other types of learning may be expected to bring to light curves of very different shape. This is particularly true of forms of learning which depend chiefly upon analysis and selection or in cases where there is no physiological or absolute limit within ordinary attainable bounds, such as, for example, learning facts of history.

Two rather extensive studies on analytical types of learning are now available. They reveal a very characteristic type of curve. The first investigation, by Ruger ('10), was based upon the number of successive solutions of a given mechanical puzzle that could be

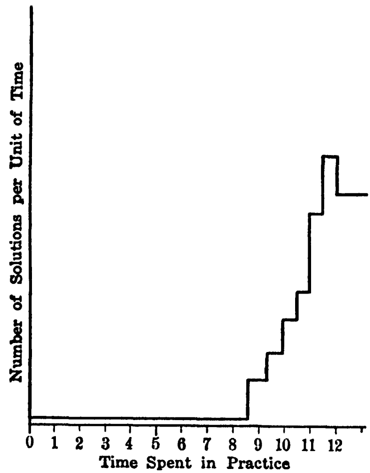

FIG. 44.—Curve to show the progress in solving puzzles. After Ruger ('10). as reconstructed by Thorndike ('14), III, p. 342.

performed by an individual within a certain period of time, Figure 44. It yields a strikingly concave curve in marked contrast to those we have previously examined. Unfortunately, practice was not continued long enough to reveal the complete curve of this type of learning. A second study, by Hull, was based upon the rate of evolution of abstract ideas as shown by their increase in ability to function, Figure 45. The material used was an elaborate system of Chinese characters combined with nonsense syllables. In this case the work was carried to the point of perfection. He found as a consequence not only the initial concave section shown by Ruger but a later period of diminishing returns. Taken alone, this last section strikingly resembles the learning curves of the simpler processes and its course is doubtless determined by the same causes. The initial plateau or period of slow progress is probably due to the necessity of making a preliminary analysis of the material used before proceeding with the remainder of the process. Clearly the elements common to many situations, as in Hull's investigation, for example, must be perceived as separate elements before they can be perceived as common elements.

It seems also quite likely that in learning facts of history or facts of science, in which there is no physiological limit and in which

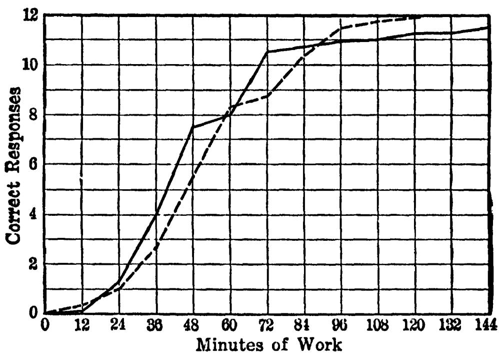

FIG. 45.—Progress of two individuals in generalizing abstractions or forming concepts. After Hull ('19).

the number of items that may be learned is practically unlimited, the course of the curve, at least for a considerable distance, is concave. This is hinted at in the results obtained with the author's tests in geography and history. Thus in the former test, given at the end of the school year to some 1,300 pupils, and in the latter test given to some 2,000 pupils, the average scores for the ends of the respective years were as follows:

Grade	5	6	7	8
Geography	25	46	72	
American History		7	20	38

These scores substantially mean that so many geographical or historical items were known to the pupil. Both sets of figures show a larger gain from the second year to the third year than from the first to the second. These data furnish of course only fragmentary portions of the curves of learning in these subjects. But so far as any form may be inferred from them, it is probably of the concave sort.

It seems probable that future experimentation will yield similar forms of curves in other types of learning. Apparently there are types of learning in which continued training brings increasing returns. Thorndike states that:

"Negative acceleration (that is rapid rise or convex form) of any great amount is far from being a general rule of learning. On the contrary, it may well be that there are some functions, such as amount of knowledge of history or geography, or of foreign languages, or of fiscal statistics, where, by any justifiable score for 'amount of knowledge' the rate of improvement in hour after hour of practice would rise, giving a pronounced positive acceleration. Each item of information may, in such cases, make the acquisition of other items easier; learning some one fact may involve knowledge of a score of new facts in the shape of its relations to the facts previously learned. So knowledge may roll up like a snowball, its sum being, say, as the cube of the amount of time spent. What we may call the 'knowledge functions' do, as a rule, show, to say the least, very much less of the diminishing returns from increasing practice than do the functions of skill in some single line of work which figure so often in the experimental studies of practice." ('14, II, p. 257.)

Whether or not plateaus occur universally in all types of learning, and whether they are really unavoidable stages in the course of learning, is an open question. They have not been found to occur as generally as the initial rise even in curves of skill. Bryan and Harter found periods of slow progress in three-fourths of their subjects. Book found them in two of his three persons, and Swift reports none. In the twenty curves obtained in the author's substitution test, eleven contained plateaus and nine did not. The practice periods totaled 120 minutes. It is possible that some of the curves might have shown plateaus if the practice had been continued longer.

Batson, who undertook an investigation for the purpose of studying plateaus, found none in the ball-tossing curves, although the training was continued for a long time, but found a pronounced plateau in learning to throw shot into a pocket.

Plateaus may be caused by lagging in energy, by loss of attention, interest, and effort, by fatigue, by periods of mechanization, and the like. Rapid progress after a plateau may be due to a recuperation in physical energy, in attention, interest, and effort, to the acquisition of new methods of learning and doing the task concerned, and to better use of the bonds which have been made automatic by the preceding practice. Bryan and Harter believe that the plateaus in the learning of telegraphy were due to the establishment of a hierarchy of habits. During the initial period of progress, the simple elements such as the signals for letters, were acquired first and when these all had been learned, there came a dead level during which the connections became automatized, and then rapid progress was again possible by virtue of the acquisition of combinations of letters into words and words into phrases. Their own statement follows:

"A hierarchy of habits may be described in this way: (1) There is a certain number of habits which are elementary constituents of all other habits within the hierarchy. (2) There are habits of a higher order which, embracing the lower as elements, are themselves in turn elements of higher habits, and so on. (3) A habit of any order, when thoroughly acquired, has physiological and, if conscious, psychological unity. The habits of lower order which are its elements tend to lose themselves in it, and it tends to lose itself in the habits of higher order when it appears as an element therein.

" 2. The Order of Learning the Habits of the Telegraphic Language.

" The synchronous curves of Figure 38 and the experience of operators agree in showing that from an early period letter, word, and higher habits make gains (a) simultaneously, but (b) not equally.

" (a) The simultaneity in these gains is shown in Figure 38 by the fact that from the point where the curves diverge, each continues to rise. This is perhaps to be explained by the fact that from an early stage the learner practises with sentences, taking them as slowly as necessary. In this way there is incidental practice of every language unit and of every language unit in its proper setting.

" (b) The curves of Figure 38 show also, however, that for many months the chief gain is in the letter and word habits, that the rate of receiving sentences, is in this period, mainly determined by the rate of receiving letters and words, and that rapid gain in the higher language habits does not begin until letter and word habits are well fixed. This objective result is supported by the introspective evidence of operators. In the first days one is forced to attend to letters. In the first months one is forced to attend to words. If the learner essays a freedom for which he is unfit, suddenly a letter or word which is unfamiliar explodes in his ears and leaves him wrecked. He has no useful freedom for higher language units which he has not earned by making the lower ones automatic. The rank and file of operators are slaves to the machinery of the telegraphic language. They must copy close. They cannot attend much to the sense of the message as it comes, but must get its form, and re-read for the sense. Only when all the necessary habits, high and low, have become automatic, does one rise into the freedom and speed of the expert.

"3. The Plateaus.

"We are now prepared to offer an explanation for the salient peculiarity of the receiving curve—its plateaus.

"A plateau in the curve means that the lower-order habits are approaching their maximum development, but are not yet sufficiently automatic to leave the attention free to attack the higher-order habits. The length of the plateau is a measure of the difficulty of making the lower-order habits sufficiently automatic."

The explanation of plateaus probably depends upon the nature of the learning process in which they occur. The theory of the hierarchy of habits would probably not apply to such a task as mirror tracing.

Experimenters are divided in their opinions concerning the inevitableness or the usefulness of plateaus even in those types of learning in which they frequently occur. Bryan and Harter, Swift and others believe that they serve a beneficial purpose. Swift, for example, says:

"The real advance in the early stages of learning is made during the periods of seeming arrest of progress. The manifest advance, that which is revealed by the curve or by examination marks, which is the same thing, is discouragingly brief. By far the greater part of the learning period is spent on plateaus when both teacher and pupil, failing to understand the situation, feel that they are marking time. Yet it is during these days of retardation that the valuable and solid acquisitions are being made. Americans who spend several years in Germany pass through a long period of discouragement. Though they study the language faithfully, and avail themselves of every opportunity to practice conversation, they seem to make absolutely no progress. The length of this plateau-period varies with different persons, but all experience its oppressiveness. Now the most curious feature of this plateau, aside from its overpowering monotony, is the suddenness with which it finally disappears. Several have told the writer that they went to sleep one night unable to understand anything, as it seemed to them, and utterly discouraged, and awoke the following morning to find that they had mastered the language, that they could understand practically everything that was said to them. The word associations and national peculiarities of thought sequence had been automatized during the long period when no visible progress was being made." ('06, pp. 310 f.)

Other investigators believe that plateaus are not necessary stages in the course of learning, but that they are due to causes which may be avoided by introducing new stimuli or new methods of attack in learning so that continued progress may be possible.

Plateaus are apparently not universal in all types of learning, nor are they found in all persons in the same type of learning. Whether they are useful stages in the learning process is a moot question. If they are not necessary, it would be highly important for education to prevent their occurrence in the learning of school material (1) by removing the conditions which bring them about, and (2) by providing stimuli at the points at which they are apt to occur so as to continue upward in the course of learning. Further experimentation will have to be made to furnish a definite solution of the problem.

Factors Affecting Progress. *a. Length and Distribution of the Periods of Work.* How long at a time, and how often, should the learner work at his task in order to make the maximum progress for the time devoted to it? Every type of learning probably has an optimum length and frequency of periods of practice. Ebbinghaus ('85), in his pioneer study of memory, found that it was better in learning nonsense syllables to distribute a given amount of time over three days than to spend it all on one day. Sixty-eight repetitions made in immediate succession were not as advantageous for later relearning as thirty-eight repetitions distributed over three days. Practically all investigators who have touched upon this phase of learning have found a principle of similar nature. Jost ('97), also working with syllables, found, for example, that two repetitions a day for twelve days were better than four repetitions a day for three days. Some of the results of both Ebbinghaus and Jost imply that in some instances a decreasing amount of time on successive days would be more economical than an equal amount on all days; that instead of distributing 24 repetitions by having four on six successive days, it would be better to have eight on the first day, six on the second, four on the third, three on the fourth, two on the fifth, and one on the sixth day.

Lueba and Hyde ('05), in an experiment on learning to transcribe English words into German script, found that of four plans of distributing time, twenty minutes twice a day yielded the slowest gain, while twenty minutes every third day yielded better, and twenty minutes every day or every other day yielded the best results.

Miss Munn ('09) made an investigation of practice in a substitution test consisting of transcribing 4,000 letters into other letters according to a key. Her distributions of time and results are given in the following table:

TABLE 38

Practice in substituting letters for other letters according to a key. After Munn ('09)

No. of Subjects	Distribution of Work	Average Time (Seconds) First 200 Letters Approximate	Last 200 Letters
23	200 letters a day for 20 successive days	41 5	13 4
4	800 letters a day for 5 successive days (400 in a. m., 400 in p. m.)	57.5	17.1
4	1000 letters a day for 4 successive days	47.00	16 5
4	2000 letters a day, seven days apart	39 5	18 2
	4000 letters in one day (1000 at a sitting)	38 5	18 5
4	3000 letters a day (at one sitting)	44.00	21.1

The highest degree of efficiency was reached by the 20-day group who reduced their time for the last 200 letters to 13.4 seconds. A definite comparison is a little difficult to make owing to the large differences in initial ability among the various groups.

In the substitution test carried out by the author, Figure 46, ten minutes twice a day was productive of the greatest progress, twenty minutes once a day was productive of almost as rapid progress, forty minutes once a day was productive of considerably less progress, while 120 minutes at one time produced scarcely half as much progress as the ten-minute or twenty-minute periods. The total time in all four distributions was the same.

Dearborn ('10), who reported an earlier experiment with the same substitution test, divided the subjects into two groups working ten minutes once a day and ten minutes twice a day respectively. He found a small advantage in favor of the former group.

Pyle ('13), working with a substitution test, reports that:

"Generally speaking, daily practice seems to give better returns than the same number of periods distributed on alternate days or in twice-a-day periods. However, there is some evidence that in the early stages of habituation, the second practice on the same day gives good returns and that, later on, alternate days may be the best distribution."

Kirby ('13) carried out a practice experiment in addition and division with 1,300 pupils in the third and fourth grades. The pupils practiced addition in 22.5, 15, 6, and 2-minute periods, and division in 20, 10, and 2-minute periods with the following gains:

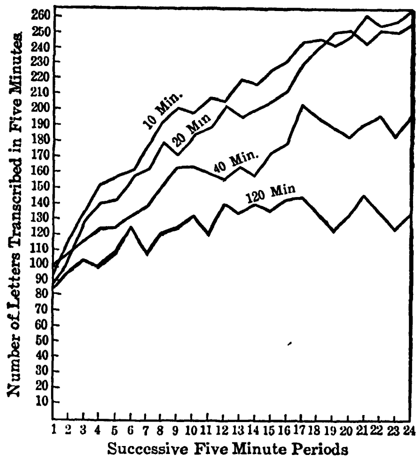

FIG. 46.—Practice in writing numbers for letters according to a key. After Starch ('12).

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

TABLE 39

Period	Addition Per Cent Gain over the 22.5-Minute Period	Period	Division Per Cent Gain over the 20-Minute Period
22.5	....	20	......
15.	21 %	10	10.5%
6.	1 %	2	77 %
2.	46.5%		

The superiority of the 2-minute period is probably exaggerated, as Thorndike has suggested, by the greater opportunity for outside practice and longer continuation of regular school work, since this period was extended over a larger number of days.

Thorndike ('11) compared the improvement in multiplying mentally with three-place numbers continuously, for two to twelve hours with Miss Whitley's ('11) subjects who did three similar problems a day for twenty days. The outcome was in favor of the distributed practice, but probably only slightly so when allowance is made for the effect of fatigue in the continuous work.

In general, relatively short periods of work in simple associative learning are probably the most economical. It would be unwise in the absence of more extensive experimental studies, to generalize to all types of learning and particularly to the learning of school subjects. What the most productive periods for learning reading or spelling or Latin or English composition are, will have to be determined experimentally in each case. All that we can say at present is that each type of learning probably has its optimum length and distribution of practice periods. Lyon has stated the matter in the following manner as a result of his studies on this problem:

"With reference to the problem of the most favorable distribution of single reading. . . . I would say that the most general statement that can be made, taking all materials and methods of presentation into consideration, is that the most economical method is to distribute the readings over a rather lengthy period, the intervals between the readings being in arithmetical proportion. For example, with one individual, in memorizing a poem of twenty stanzas, the highest retentiveness was obtained by distributing the readings as follows: two hours, eight hours, one day, two days, four days, eight days, sixteen days, thirty-two days, etc. The practical bearing of the results obtained on education in general is that when associations have once been formed, they should be recalled before an interval so long has elapsed that the original associations have lost their color and cannot be recalled in the same shape, time, and order. In general it was found that the most economical method for keeping material once memorized from disappearing was to review the material whenever it started to fade. Here also the intervals were found to be, roughly speaking, in arithmetical proportion. For similar reasons the student is advised to review his lecture notes shortly after taking them, and, if possible, to review them again the evening of the same day. Then the lapse of a week or two does not make nearly so much difference. When once he has forgotten so much that the various associations originally made have vanished, a considerable portion of the material is irretrievably lost." ('13, p. 161.)

b. Forgetting. Learning is, in a certain sense, a fortification against forgetting, and from the practical side, it is important to know how frequently and in what manner the fortifications should be strengthened in order to resist the attacks of forgetting. Only a few experimental studies have been made on the rate and factors of forgetting. Ebbinghaus ('85) learned nonsense syllables until he could give them once correctly, and then measured the rate of forgetting by the amount of time required for relearning them at different intervals after the original memorizing. Radossawljewitch ('07) used nonsense syllables and poetry, Bean ('12) used series of letters, and Magneff [1] used poetry. The curves of for-

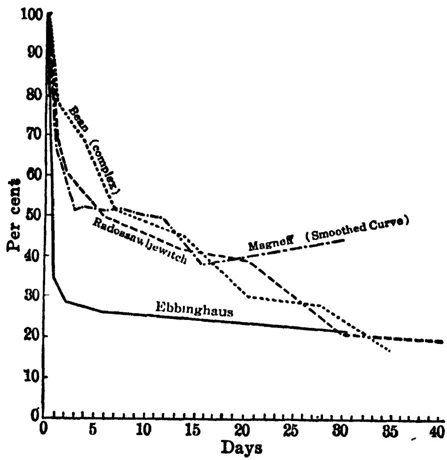

FIG. 47.—Curves of forgetting.

getting obtained by these investigators are presented in Figure 47. They agree in showing a very rapid rate of loss at first, followed by a very gradual decline afterwards. In retaining syllables, Ebbinghaus found that he forgot as much in the first twenty minutes as in the following thirty days; in remembering a poem which had been learned to the point of two perfect repetitions, Radossawljewitch found that his subjects forgot as much in the first two days as in the next twenty-five days.

Inquiry into the rate of deterioration of connections through lack of practice have also been made by Book ('08), Rejall,[2] Swift

[1] As reported by Radossawljewitch.
[2] Reported by Thorndike, II, p. 309.

and Schuyler ('07) in the case of typewriting, and by Swift in the case of tossing balls. These seem to indicate much greater permanency in sensori-motor connections than in the memory of syllables or poetry.

The experimental work on forgetting is too limited as yet to permit of much definite application in the practical procedure of learning school material. The one suggestion that may possibly be made would be this: Since the rate of forgetting is very rapid at first and more gradual later on, it probably would be highly advantageous to have relearning of a given material come very frequently at first and more rarely later on. Thus the facts of a lesson in history or the newly acquired words of a spelling lesson should be reviewed the next day or perhaps preferably on the same day, then again two or three days later and then a week or ten days later, and so on.

The effect of long vacations upon the retention of school material has been investigated only partially. Measurements of skill in arithmetical operations in June and September show heavy losses (see page 403) and raise the question as to whether long vacations are really profitable or detrimental.

c. Concentration, Effort, and Zeal. "Practice makes perfect," is only a half truth. Only practice with zeal and effort is likely to bring improvement. A great deal of practice and repetition may continue day in and day out without the slightest gain. While the factor of zealous attention and interest has long been recognized as a matter of common-sense observation, its real value, however, has never been appreciated until experimental studies pointed out its actual significance. Bryan and Harter have called attention to this point in a very emphatic manner as follows:

"A fact which seems to be highly significant is that years of daily practice in receiving at ordinary rates will not bring a man to his own maximum ability to receive. The proof of this fact is that men whose receiving curve has been upon a level for years, frequently rise to a far higher rate when forced to do so in order to secure and hold a position requiring the higher skill. That daily practice in receiving will not assure improvement is further seen in the fact that in many cases inferior operators after being tolerated for years are finally dropped because they do not get far enough above the dead line. One conclusion seems to stand out from all these facts more clearly than anything else, namely, that in learning to interpret the telegraphic language, it is intense effort which educates. This seems to be true throughout the whole length of the curve. Every step in advance seems to cost as much as the former. Indeed, each new step seems to cost more than the former. Inquiry at the telegraph school and among operators indicates that between sixty and seventy-five per cent of those who begin the study of telegraphy become discouraged upon the plateau of the curve just below the main-line rate. As a rule, ordinary operators will not make the painful effort necessary to become experts. Facts of an analogous character will be recalled from other fields.

"The physiological, psychological and pedagogical implications of this conclusion are manifestly important. If in our educational methods in the past, we have often made the pace that kills, there is possibly the danger on the other hand that we shall make school work all play, and so eliminate the intense effort which is necessary for progress." ('97, p. 50.)

A great deal of learning is done without any real zeal or effort toward improvement. The usual way in which a great many children learn to play the piano illustrates how much practicing and learning consists in dawdling with more attention upon the clock than upon the music sheet. A great deal of learning of school material is done with the same lack of interest and effort.

d. Specific Practice in the Functions to be Improved. One of the striking discoveries of experimental investigations is the very rapid progress in specific functions when the practicing is done on the particular connections to be established. A surprisingly large percentage of pupils make little or no progress in an entire year's work in such subjects as reading, writing, and the like, while the remaining pupils make only a part of the progress that they could make if their efforts were squarely directed at the material to be learned or at the associations to be established.

The numerous practice experiments that have been conducted in psychological laboratories during the last two decades furnish overwhelming evidence of the tremendous improvement obtained under experimental conditions. Only a few examples will be cited.

The writer found that eight persons, practicing mental multiplication of three-place numbers by one place-numbers for about 15 minutes a day for 14 successive days, made enormously large gains as shown in the following table:

TABLE 40

Improvement in mental multiplication. After Starch ('11)

Individual	Examples Done per 10 Minutes on 1st Day	Examples Done per 10 Minutes on 14th Day	Gross Gain	Percentage Gain
S.	25	62.5	37.5	150
St.	37.7	81	43.3	115
F.	23.8	45.4	21.6	91
V.	41.7	71.4	29.7	71
W.	14.7	29	14.3	97
H.	37	100	63	170
Si.	25	29.8	4.8	19
B.	23.4	66	42.6	182

These subjects gained in approximately four hours of practice per person over 100%, varying, of course, from one person who made little gain up to two persons who gained nearly 200%.

Wells found the amounts of gain from 150 minutes of practice in addition on the part of ten adults as follows:

TABLE 41

Improvement in addition: adults. After Wells ('12)

Individual and Sex	Number of Additions in Five Minutes			Percentage which Amount Done on 30th Day was of Amount Done on 1st Day
	First Day	30th Day	Gross Gain	
1 f.	150	280	130	187
2 m.	180	380	200	211
3 m.	200	430	230	215
4 f.	220	380	160	173
5 m.	225	368	143	164
6 m.	225	460	235	204
7 f.	235	570	335	243
8 f.	250	440	190	176
9 f.	260	540	280	208
10 m.	290	540	250	186

The gains show approximately a doubling in efficiency in the course of thirty days.

Dearborn used vocabularies and poetry in learning experiments and reported the following results:

TABLE 42

Improvement in ability to memorize. After Dearborn ('10)

Subject	Total Practice Time in Hours	Amount Learned Daily	Number of Days of Practice	Total Amount Learned	Time Required on First Day	Time Required on Most Efficient Day
Learning the English meanings of French or German words:						
1	6 1/3	50	21	1050	30	13
2	6	35	20	700	30	12
3	6	35	18	630	30	14
4	8 1/10	30	22	660	33	15
5	7 2/3	30	20	600	40	15
Learning poetry:						
7	3 1/3	32	15	480	38	7
8	3 2/3	18	16	288	30	8
10	4 [1]	17	13	221	30	12

Similar results have been reported by Bair in tossing shot, by Swift and Batson in tossing balls, by Whitley in drawing lines in a maze, by Wells in tapping, by Kline, Wells, and Whitley in cancellation tests, by Thorndike, Wells, and Kirby in adding, by Swift, Book, Rejall in typewriting, by Ebert and Meumann, Winch, Sleight, Dearborn, and Fracker in memory—in fact in all experimental work in which practice enters. Improvement of mental functions through practice is well-nigh universal and the amount of improvement through specific training under experimental conditions is almost incredible, particularly when we contrast with it the gains made in school functions in from 50 to 150 hours devoted to a subject in the course of a year.

The average gain made by pupils in school activities in the course of a year's practice as indicated by the standard scores derived from measurements with tests and scales is shown in the following percentages of gain at the end of the eighth grade as compared with the end of the seventh grade:

[1] Approximate.

TABLE 43

Based upon published scores for the various tests

	7TH	8TH	PER CENT GAIN
Reading: speed, words per second	3.6	4.0	11
Reading: comprehension, words written	45 0	50 0	11
Writing: speed, letters per minute	75.0	83.0	11
Writing: quality, Thorndike scale	10.4	10 9	5
Addition: Courtis Series B—rights	6.5	8.0	23
Subtraction: " " " "	8.5	10.0	17
Multiplication: " " " "	6.5	8.0	23
Division: " " " "	7.0	9.0	28
Reasoning: Starch Arith. Scale A	11.0	12.6	15
Language: Starch Gram. Scale A	8.0	8.3	4
Composition: Hillegas Scale	41.0	46.0	12

These gains are surprisingly trivial when compared with the gains, often running over 100%, reported in connection with experimental investigations of practice.

Definite experimental results are not at hand to substantiate the following assertion, which may seem doubtful but which is not impossible from the present indication of other measurements, such as those presented by Dearborn in Table 42 or by the writer in Table 9 (Experiments), namely, that the average high school pupil could learn in 20 minutes a day for thirty days, all the Latin words (500 words) that he would need in an entire year of Latin. He could learn in 30 minutes a day for one-half the school year, all the Latin words (approximately 2,000) that he would use in his entire study of four years of high school Latin.

The difficulty with the material of school subjects is that we do not, and in some instances we cannot, specify with sufficient definiteness just wherein the improvement is to be made. We can point out specifically to a child whether or not he spells a word correctly and what part of the word may be incorrect, but we have not until recent years made any attempt at determining which particular words a child really ought to know how to spell. The pupil was given a list of words selected more or less on the basis of their unusualness and difficulty rather than upon the basis of usefulness or frequency of occurrence. The idea seemed to be that if he learns to spell a sufficient number of difficult and useless words, he will know how to spell all other words in the English language. The school has virtually said, "Learn to spell," but has not said what a child should learn to spell. Even in a subject in which the associative bonds may be precisely defined so that they can be directly attended to, we have not done so. The same situation obtains in practically all other subjects with the added difficulty that in some subjects the material is of such a nature that specific directions and specific material or specific bonds to be formed, cannot easily be isolated. This is particularly true of such a subject as English composition which represents the opposite extreme from spelling and arithmetic. The child is told to improve his style, or his language, or his expression, or his originality, or his imagination; but he is not told very definitely how he may do this, or just what he is to do. The school should, therefore, aim to specify exactly what sort of learning is to be done so that a definite notion on the part of the learner may be formed of the precise bonds and connections to be made.

e. Definite Knowledge of Success and Error. Much experimental work implies that the feeling of satisfaction resulting from successful trials of a task facilitates the formation of the connections concerned. It seems obvious therefore as a practical matter that precise knowledge of the success or failure on the part of the learner is exceedingly important. It will not only serve to arouse the feeling of satisfaction but also help to define the particular bonds to be established. Feelings of dislike on the part of the learner toward the material to be learned undoubtedly interfere with the rapid formation of the connections, and frequently the feeling of dislike is accompanied by an attitude of unwillingness or stubbornness indicated by such statements as "I know I can't learn languages; I never could." "I never was able to get mathematics." "I can't memorize anything." A concrete case that came under the writer's observation was that of a man considerably older than the average university student, who in the experiment on the transference of training (Chapter XI, *Experiments in Educational Psychology*) reported that he was unable to learn vocabulary and that the net result of half an hour's work on the first list was ten words. The average student is able to learn the entire list of thirty words in from twelve to fifteen minutes. He further stated that he had always had great difficulty in learning languages. In order to ascertain, if possible, the real status of his memory and other abilities, he was tested by Terman's revision and some additional tests, all of which indicated that he was of average intelligence and that his memory was not defective, but approximately average.

He was informed of the results of the tests, that his defective memory was largely illusory, and that probably his real trouble lay in his contrary attitude toward certain tasks, which was also indicated by his own statements concerning his work. The general effect upon his later attitude in learning was a distinctly wholesome one. This case is cited because it exemplifies many similar instances of persons who *feel* incapable of learning certain things.

f. Interest in Improvement. One important element in the remarkably large amount of gain through practice in specific functions, is the fact that the progress is directly observable and definitely measureable which in turn produces a real zeal toward improvement and toward outstripping the preceding records. In an experiment such as the substitution test or the practice experiments in arithmetical operations, the observation of a definite gain is possible so that the learner can see just how much progress he is making. The practical value of this suggestion would be the creation of circumstances for the learning of school subjects similar to the conditions of learning in laboratory experiments by introducing forms of measurement through which the progress may be determined at frequent intervals, so that the pupil may see what progress he is actually making.

In a certain elementary school a series of standard tests [1] was applied every month throughout the entire school year. Tests in reading, writing, spelling, and arithmetic were given at monthly intervals to determine the progress made. Each pupil knew his own record in each test and compared it from month to month. This condition developed a remarkable interest and zeal in striving to surpass the record of the preceding month. The condition created thereby was similar to that of a learning experiment under laboratory conditions. Each pupil kept his own score and knew what gain he was making each month. This condition had a remarkable effect upon the total progress made during the school year as shown in the accompanying graphs. These results show that the pupils made on the average a gain in some studies twice as great as that made ordinarily in the course of a school year. This gain cannot be attributed to familiarity with the test material since, in the case of reading, different passages were used in each successive test; in writing, a different sentence was used

[1] The tests were made in The Alice School, Hibbing, Minnesota, by Principal L. J. Coubal, and reported in an unpublished thesis in the library of the University of Wisconsin. The tests were carried out under the direction of Professor V. A. C. Henmon.

each time; in spelling, the author's six lists were used in rotation, one at a time; and in arithmetic the three sets in the Courtis Series A were used and rotated so that there was a recurrence of the same

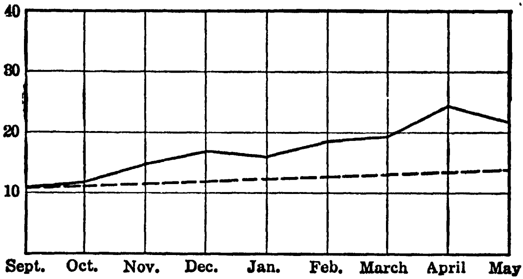

FIG. 48.—Progress in speed and comprehension of reading combined into single scores as measured by monthly tests (Starch reading tests) upon 4th grade pupils. The continuous curve represents the progress of the class. The straight, broken line is the progress for schools generally based upon the standard scores for June of the 3rd grade and June of the 4th grade.

test every three months, but it is very unlikely that this contributed any appreciable share toward the gain shown. It would seem, therefore, highly desirable if there could be introduced into the schoolroom a similar atmosphere of motivation such as obtains

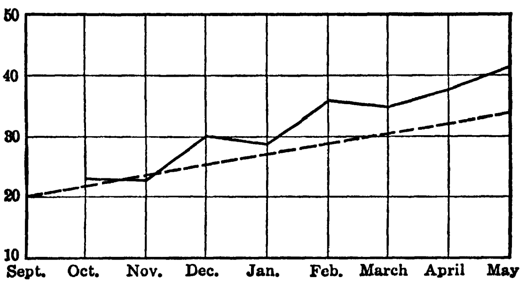

FIG. 49.—Progress in speed and quality of writing (Thorndike scale). Other facts same as for Fig. 48.

in learning experiments in the laboratory. The knowledge of one's actual ability and of the actual amount of gain serves as an exceedingly powerful spur for the learner to surpass his own previous performances. The popular dictum "Nothing succeeds like success" may be partly justified by such experimental results as the ones here cited. To see oneself gaining tends to stimulate greater efforts toward gain. The educational measuring scales and tests

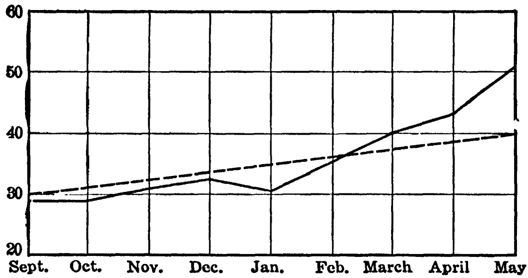

FIG. 50.—Progress in spelling (Starch test) of 3rd grade class. Other facts the same as for Fig. 48.

ought to serve a useful purpose at this juncture. They will provide means whereby the pupil may be able to see for himself in definite terms the gains he is making.

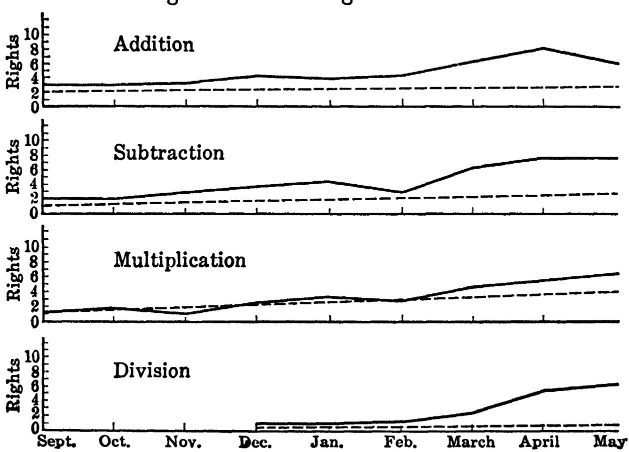

FIG. 51.—Progress in the fundamental operations in arithmetic (Courtis tests). Other data same as for Fig. 48.

g. Mental Imagery. After the early studies on mental imagery became known, there followed considerable theorizing as to types of persons and types of learners, and with it came the resulting endeavors to make applications of these theories to methods of teaching. Thus it was said that if a pupil has difficulty in learning to spell or in learning a foreign language he may be devoid of, or weak in visual imagery; or if he has trouble in learning to write, he may be short on motor imagery; or if he finds it hard to learn the pronunciation of words, he may be defective in auditory and motor imagery; and if he fails in the academic subjects he was suspected of being devoid in visual and auditory images and strong in motor imagery and should therefore turn to manual training. The proposed remedy was that the material to be learned should be presented to a different sense organ so that the pupil might use the imagery natural to him. All these inferences may possibly be true; but later additions to our knowledge of mental images make us more hesitant regarding the real part played by them in learning and concerning the actual differences produced by presenting material to different sense avenues.

Before we can make changes in practice we must be sure of the principles upon which the practice is to be based. It is important, therefore, to examine at least the following three considerations:

In the first place, more careful studies of the sorts of images employed by different individuals show that the classification of persons into visuals, audiles, motiles, and so on, is fundamentally misleading. Studies by Betts ('09) and others have helped greatly to clarify the matter by showing that pure types rarely exist. During the last six years several hundred students have performed the imagery tests outlined in Chapter VII of the author's *Experiments* ('17). Among this entire number not more than two or three persons were found whose images either were practically all of one type, or who had one or another commonly prevalent class almost entirely missing. The facts for 95% of all persons are substantially as set forth on page 45 (*Experiments*), namely, that nearly all persons have all types of images which are combined in different individuals in varying proportions. Mankind as a whole does not fall into sharply or even vaguely divided groups of visuals, audiles, and so on. They are not found except in rare instances.

In the second place, recent inquiries indicate that images of the class most natural or predominant for a given person may be aroused by stimuli coming through another sense. For example, auditory stimuli may arouse visual images as well as, or even more readily than, auditory images if visual images are more natural to the individual. Miss Abbott ('09) found in a detailed investigation with four subjects, on the memory consciousness in orthography that "irrespective of the method of presentation and the manner of learning, the typical mode of recall for all observers was through the visual imagery of the letters" (p. 153). Consequently it does not follow that, even if a person has a strong leaning toward one or another type of imagery, it is necessary or even advantageous to present the material to be learned through the sense avenue of the favorite type of imagery.

Colvin and Myers ('09) made an extensive series of tests with school children to ascertain to what extent visuals retain visual material best, audiles auditory material, and motiles motor elements. He concluded that:

"There seems to be a fairly definite relation between the effectiveness of memory in the case of a particular ideational type and the memory material which is most suited to that type. In particular the visual type retains best material with a visual content and the auditory and motor types, to a less degree, material with an auditory or motor content, as the case may be."

While this conclusion is probably justified, the differences among the various types of elements retained by the different groups of pupils are in most instances small.

In the third place, the most advantageous mode of sensory presentation for a given person is not necessarily the one that corresponds to his dominant type of imagery. Much of the experimental evidence that has been accumulated on the problem of modes of presentation of the learning material is conflicting in character. It is uncertain whether visual, auditory, motor, auditory-motor, or visual-auditory-motor presentations are most advantageous. Henmon ('12) has reported an experiment which has thrown important light upon the problem. He employed four methods of presentation, visual, auditory, visual-auditory, and visual-auditory-motor; and used three sorts of material, concrete nouns, two-place numbers, and nonsense syllables presented to six subjects with one, two, or three repetitions.

"In the visual presentation the subjects read the stimuli directly from the rotating drum and immediately wrote down as many members as could be recalled and in the order presented. The subjects were asked to repress movements of articulation. In the auditory presentation the experimenter read the stimuli from the drum, the subject keeping his eyes closed and repressing movements of articulation. In the visual-auditory presentation the subject both saw the stimuli and heard them read by the experimenter. In the visual-auditory-motor presentation the subject himself read the lists aloud."

His results are summarized thus:

"1. Auditory presentation is clearly superior to visual presentation in immediate memory of adults, a result attributable to the greater ease and freedom of visualization with auditory presentation and the greater effort of attention required.

"2. This superiority of auditory over visual presentation holds for all materials (nouns, nonsense-syllables, numbers), for all subjects irrespective of image type, and for one, two and three presentations. This result is not in accord with the opinion commonly held that visual presentation is superior, especially with meaningless material.

"3. Combined visual-auditory presentation is slightly inferior to the auditory alone and decidedly superior to the visual alone. The advantage of a combined method is very much less than that shown in earlier investigations.

"4. Visual-auditory-motor presentation is slightly inferior to the auditory and the visual-auditory presentations and superior to the visual alone. Articulation or vocalization is of little value for immediate memory.

"5. The correlations of abilities with different forms of presentation are positive and very high, superiority with one indicating practically the same degree of superiority with another." (After Henmon '12.)

A fair general impression of the present status of our knowledge of imagery in relation to learning would be that distinctions among types of pupils have been overemphasized and that much of the endeavor to adjust methods of teaching accordingly has been misdirected.

h. Fatigue. As a final important factor in the progress of learning, we must consider fatigue. Educational and psychological literature has been replete with discussions regarding the part which fatigue plays in the reduction of mental efficiency. While psychological research has provided considerable information concerning the course of continuous work and concerning the changes in the efficiency of the worker as measured by cross-sectional tests at various stages of work, it has not furnished as yet much definite knowledge concerning the control of the work of the pupil in school. Pedagogical literature has been generous in pointing out signs of fatigue and serious consequences of overwork and in suggesting remedies for avoiding exhaustion, yet we are not sure whether the so-called symptoms are indications of real fatigue or whether any serious or even mild fatigue effects ever result from the work as carried out in the great majority of schools.

In discussions of fatigue it is important to bear in mind two distinctions in the meaning of the term, namely, fatigue in the sense of decrease in the capacity to do work, and fatigue in the sense of decrease in interest in, or willingness to, work. The two are plainly different and do not necessarily go together. The one is actual loss in efficiency; the other is a feeling of ennui or weariness. Much of our thinking about the problem has been confused by a failure to distinguish between these two meanings. Fatigue in the former sense probably has been greatly exaggerated as an educational problem. Perhaps only in exceptional individuals is there injurious overstrain due to mental work. The discussion of this topic will, therefore, be abbreviated.

The experimental methods by which the phenomena of fatigue have been investigated will first be mentioned briefly. They may be divided into two classes: (1) Indirect methods, and (2) Direct methods.

(1) Indirect methods. The principle, upon which the indirect methods have proceeded, has been to measure some physiological or psychological functions at different points during the course of work in order to compare the efficiency of those functions on the assumption that the difference in them would be indicative of efficiency in general. One of the first methods was that employed by Griesbach, who determined the two-point threshold upon various parts of the skin at various times of the day on the belief that a decrease in sensitiveness or a widening of the threshold indicated a reduction of general mental efficiency. He made extensive comparisons among school children for the purpose of determining the amounts of fatigue produced by various types of school work, and formulated an elaborate series of conclusions with regard to them. For example, specific fatigue values were assigned by him and his followers to the different school subjects. Vannod states that mathematics, Latin, and Greek produce most fatigue, and that French and geography produce least. The difficulty, however, with results of this type is that while the two-point discrimination upon the skin varies under different mental and physical conditions, it is a rather unsafe basis upon which to make sweeping generalizations concerning the general mental efficiency of a person. In

fact, the closeness of the agreement of the size of the two-point threshold with the actual amount of fatigue is too uncertain to use this function as a symptom of general mental or physical fatigue. A number of other indirect methods have been employed, such as the rate of tapping with a stylus, the variation in blood pressure, in pulse, in respiration, the range of visual accommodation, sensitiveness to pain, and so on. The same criticism applies to these as to the two-point discrimination. These functions may have concomitant variations within rough approximations, but they are too distant to be precise indications of mental efficiency.

The use of the ergograph as developed by Mosso and his co-workers has probably been the most successful and useful method for studying problems of fatigue. As such it is, however, a direct method for investigating muscular work and fatigue and only a very indirect and doubtful method for investigating mental fatigue.

Other indirect methods of a more distinct psychological character have also been employed. These have consisted of the measurement of certain mental functions at various intervals in order to determine how much variation there may be in these functions and to regard them as indications of mental efficiency in general. Such tests have been made upon memory, various types of association processes, perception as measured by cancellation tests, and the like. These tests have a certain superiority over those mentioned in the preceding paragraph since they deal at least with psychological functions, but they likewise do not directly measure the course of work as it actually occurs. They have, however, been useful in comparing efficiency in the same mental capacities at various points during the course of a day.

A considerable number of researches by means of cross-sectional test methods have been carried out upon school children as well as adults. Thus, for example, Sikorski ('19) tested pupils before and after school in writing from dictation, and compared the number of errors made. Bolton ('02) measured the memory span for digits during the early and the later part of the school day. Laser ('94) made a test with pupils in addition and multiplication at hourly intervals. Friedrich ('97) tested 51 pupils in addition, multiplication, and in dictation exercise at hourly periods. Ebbinghaus ('97), with the aid of the teachers, gave tests at hourly intervals in immediate memory of numbers, in addition and multiplication, and in supplying words and syllables omitted from sentences. Ritter (1900) used tests in dictations of words, numbers, and sentences, and tests in cancelling letters and words. Thorndike (1900, '11 and '12) used, early and late in the school day, tests in adding, multiplying, cancelling certain words in a printed text, and memorizing numbers, letters, and geometrical forms. Heck ('13) measured the performances of pupils in adding, subtracting, multiplying and dividing at four points during the school day. Miss King [1] used tests at five points during the school day in adding and multiplying and in answering questions of a general informational character.

Practically all of the investigations here mentioned that were carried out reliably, agree, when interpreted fairly, in showing that efficiency in the various functions examined is changed very slightly or unappreciably during the course of a school day. Not all of the investigators, however, interpret their results in this manner. Thorndike has pointed out a very important misconception in the interpretation put by some of the experimenters upon their data, namely, that of counting simply the number of errors made at different times of the day instead of expressing efficiency in terms of both amount and accuracy of work done. This point may be illustrated in the case of Friedrich's results presented in Table 44.

TABLE 44. After Friedrich

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

[1] An unpublished study reported by Thorndike. ('14, III, p. 93).

If in this table we compare simply the percentage of errors the efficiency of the pupils was over five times as great at the beginning of the school day as at the end in the dictation test, and over two times as great in the computation tests. If, on the other hand, we consider the column giving percentage right we find that the efficiency changed but very little.

To point out further how inconsiderably the performance of pupils changes in the course of a day we may note the following results from Heck:

TABLE 45. After Heck ('13)

Time of Test	Units of Work Done	Per Cent Correct
9:10 a. m.	140.37	87.40
11:05 a. m.	142.57	86.08
1:10 p. m.	142.67	86.17
2:30 p. m.	143.68	85.46

The amount of work increased slightly while the accuracy decreased slightly from the first test to the last.

(2) Direct methods. The most fruitful direct methods of measuring continuous mental work have been the various types of mental calculations, particularly addition and multiplication. These methods have been used by Krapelin, Thorndike and Arai, Starch and Ash, and others. As an illustration of one type of mental addition, the writer has used a method consisting in the mental addition of 6, 7, 8, and 9 in rotation by beginning with a given number and adding each of these numbers in turn to the answer last obtained, as described in Chapter XVI, *Experiments* ('17). The advantage of this form of calculation is that it affords sufficient difficulty and thus fully taxes the efforts of the individual and makes possible a minute record of the amount and accuracy of work done during succeeding short intervals of time. Figure 52 shows a curve obtained by this method, covering a period of continuous work of two hours.

As an illustration of mental multiplication, we may cite the experiment carried out by Miss Arai under the direction of Thorndike. She used the method of multiplying mentally four place numbers by four place numbers, as 4,962 times 7,584. She trained herself for a considerable period of time in this type of mental multiplication in order to reach an approximate limit of practice. Then she did the following experiment:

"On March 3, 4, 5, and 6, that 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 remarkable result of all experiments with purely mental functions has been that mental efficiency is reduced only very slightly even after two or more hours of very difficult, uninterrupted

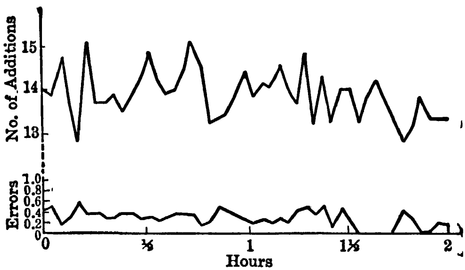

FIG. 52.—Mental work curve. Upper curve shows number of additions made per half minute period. Lower curve shows number of errors made. Work was continued for two hours. After Starch and Ash ('17).

work. Thus in the curve, Figure 52, the reduction in the number of additions made per thirty seconds, was only from 14.0 down to 13.4, or a loss of only 4.3%. Arai found even in the course of 12 hours of such difficult mental multiplication as she carried out, that her efficiency was reduced only by about one-half. Other investigators have shown in general the same facts.

Seashore and Kent ('05) measured continuously, for as long as two hours, the threshold of hearing by recording the audibility and inaudibility of a sound varied about the limen. The intensity of the sound was changed at a uniform rate. As soon as it became too faint to be heard the subject gave a signal to the experimenter who at once increased the strength of the sound. As soon as it could be heard again the subject again responded. Then the sound was decreased again, and so on without break. A sample curve is shown in Figure 53. Ten records were obtained which showed that "continuous liminal or moderately faint sounds do not seem to lower the efficiency of the ear in a two hour test" (p. 100).

It would seem, therefore, on the basis of experimental work, that fatigue in the sense of decrease in product achieved is practically a negligible element in school work. The actual capacity to do work with the same degree of accuracy is practically undiminished in the course of a school day. Such symptoms of fatigue as have been frequently enumerated in pedagogical writings, are apparently only superficial signs of monotony, of lack or diminishing of interest, or of being bored by school work, and not actual signs of loss of capacity to do the work. Such statements as "I simply cannot work any longer" made after a half or whole hour's work, are illusory and probably signify chiefly a weariness with the work

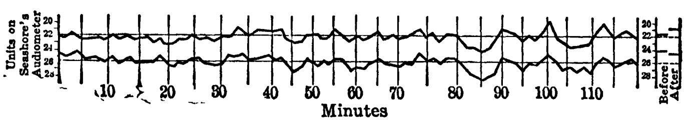

FIG. 53.—Continuous record of the measurement of the threshold of hearing. After Seashore and Kent ('05).

which, if it must be kept up by force of conditions, can usually be continued without difficulty or harm and usually without being seriously boresome.

The feeling of interest or satisfaction in doing work does decrease very materially as the work goes on. Thorndike (17) for example found that the satisfyingness of such work as grading compositions decreased in the course of two hours to about one-half and in the course of four hours to about one-third of the amount of satisfaction at the beginning of the period.

The feeling of weariness, from the practical side of school activities as well as of mental work generally, is, however, an important item. In a certain sense it is a real thing. Even if it is illusory it does interfere with the smooth continuation of work. But it is very likely a less serious situation than an actual loss of capacity to do work would be. Practically it resolves itself into a problem of maintaining interest rather than relieving depreciation of efficiency.

CHAPTER XII

HOW TO STUDY

Waste in Studying. Since studying is learning under school conditions, it would seem worth while to make such suggestions as can be made concretely to assist pupils in this important phase of the psychology of learning. It may seem preposterous to give advice about something concerning which each pupil is presumably proficient after years of practice in it, and furthermore to attempt to give suggestions on studying may seem to many to be nothing more than an "unprofitable delineation of the obvious." It is, however, very certain that there is an uncalculated waste of energy and a still more prodigal waste of time in so-called studying. If we may judge from the possibility of improvement in reading capacity alone, and from the larger accomplishments attained under favorable conditions of work, we may venture to guess that the average student could accomplish his work just as efficiently or more efficiently, in two-thirds, or less, of the amount of time ordinarily consumed, by developing more economical methods and habits of studying. Improvements in proper procedure in studying have shown how much more may be accomplished in the same length of time or even in a shorter period of time. Vicious habits of dawdling in school work are acquired, which may have their permanent effect throughout the individual's life.

Is Studying Worth While? This question is worth raising in view of the belief, prevalent among students, parents, and graduates, that after all it does not matter much whether a pupil does well in his studies or not, that the boy who does poorly in the grammar grades or the high school will outgrow his negligence and come into his own when he gets into his college or professional course, or that when he gets into the real business of life he will outstrip his more studious mates. To what extent are these beliefs true or false? To what extent is early scholastic performance indicative of similar or different performance later on? To what extent is scholastic performance prophetic of performance in life?

A considerable amount of statistical material has been accumulated in the attempt to answer these questions. Some of this material was presented in the latter part of Chapter IV under the heading "Correlations between Early and Later Mental Abilities." These correlations were found to be high. Dearborn, for example, found that of 472 pupils, whose records were traced through the high school and college, only two who were in the lowest quarter in the high school rose to the top quarter in the university. It should be noted further that these two were just barely poor enough in the high school to be classed into the lowest quarter and that they rose just barely enough to get into the top quarter in the university. The chances that the pupil who is doing poor work in the high school will later come into his own are exceedingly small; apparently he has been in his own all along, or, if not, he had better have got into his own as soon as possible.

President Lowell ('10) made a study of the records of the graduates of Harvard College for a period of twelve years. He found the following situation:

Men graduating with various honors	Percentage graduating with distinction from The Law School	The Medical School
A. B.'s with highest honors......	60	92
A. B.'s with great honor.........	40	87
A. B.'s with honor..............	22	76
A. B.'s without honor...........	6½	36
A. B.'s without honor, of men who had entered college with conditions......................	3	

The 250 Yale men who graduated from the Harvard Law School in 1900–1915 were divided into nine groups according to their scholarship at Yale. These nine groups, with the exception of one, finished the Harvard Law School in the same relative order of scholarship that they had held at Yale.

To many persons a more important problem is the relationship between scholastic attainment and success in business or professional work. Foster ('16) has summarized in an interesting manner much of the evidence pertaining to this problem. He made a study of the Harvard College class of 1894. He asked three men, the dean of the college, the secretary of the alumni association, and a member of the class, to name the most successful men of the class. They were free to use their own interpretation of success except that they were not to include men whose success appeared to be due chiefly to family wealth or position. The three judges agreed on twenty-three men. Foster then obtained their records in college and compared them with the records of twenty-three other men chosen at random from the same class. The former had nearly four times as many highest grades as the latter, namely, 196 A's as compared with 56 A's. By a similar plan three judges selected the most successful men among the graduates of the University of Oregon for the period of 1878 to 1901. Of the graduates designated as successful, 53% had been good students and 17% had been weak students. Of the graduates designated as unsuccessful, 12% had been good students and 52% had been weak students.

A study of the alumni of Wesleyan University showed that of the living graduates for the period of 1860 to 1889, 50% of the men who had graduated with honors were listed in *Who's Who*, and only 10% of the men who had graduated without honors were in *Who's Who*. Among the living graduates for the period of 1890 to 1899, 60% of the men graduated with highest honors were listed in *Who's Who*, 30% of the men elected to Phi Beta Kappa were listed in *Who's Who*, while only 11% of the graduates without superior scholarship were found in *Who's Who*. (Nicholson '15).

E. G. Dexter investigated the records of the living graduates of twenty-two colleges and found that 5.9% of the honor scholars and only 2% of all graduates were listed in *Who's Who*. Furthermore, 56% of the Yale valedictorians were found in *Who's Who*. Their chances were, therefore, more than twenty-five times as great as those of other graduates. The records of 13,705 living graduates of two New England colleges revealed the fact that 5.4% of those who constituted the highest tenth were listed in *Who's Who* while but 1.8% of those in the fourth tenth were there listed. *Who's Who* is, of course, not an absolute criterion of success; it is, however, a rough measure of success.

A tabulation of the Oxford University men who entered the law or the ministry showed the following percentages of men who attained distinction in their respective professions:

Men with varying honors	Percentages attaining distinction	
	In the law	In the ministry
Men with 1st class honors	46%	68%
" " 2nd " "	33	37
" " 3rd " "	22	32
" " 4th " "	20	29
" " pass degrees	16	21
" " no degrees	15	9

As a matter of correct interpretation of these extensive statistics, it must not be assumed that success or failure is solely attributable to the amount of devotion to school studies. The uniform manner with which the early scholastic records agree with the later records of the same persons, or the pronounced tendency with which scholastic attainment correlates with business or professional attainment is probably due to a common cause, namely, original ability or make-up of the individual. At their face value, these figures mean that the person who does well in his school work also tends rather strongly to be a person who will do well in his business or professional work. However, this array of facts is impressive and ought to be brought emphatically to the attention of high school and college students. They ought to have a tonic effect upon their efforts. While our native make-up determines to a large extent what we shall become, yet rarely does any one utilize or develop to the fullest extent even the limited measure of ability that he possesses. The laggard can find little consolation in the hope of somehow redeeming himself later on.

Types of Studying. For the sake of convenience, we may divide studying into three types:

1. The Reading Type of Studying. In the elementary school probably eight-tenths and in the high school and the university probably two-thirds of all studying consists essentially in reading.

2. The Laboratory Type of Studying. This type obviously consists of the manipulation of apparatus, the observation of material, the recording of observations and experimental data, and the interpretation of these data.

3. The Analytical or Reasoning Type of Studying. This type is involved in those subjects in which the amount of reading is relatively little, but in which the task consists in a thorough mastery of a relatively small amount of text. Such studying is obviously involved in mathematics and in a few other types of difficult reading, as for example, certain branches of philosophy and the speculative and theoretical aspects of the sciences.

Problems. Every type of studying is different and, in a sense, every lesson has its own special material and presents its own problems on how to study effectively. It may seem futile to attempt to give *general* advice on how to study. Yet upon further analysis, it appears that there are several elements common to all types of learning. These elements are (1) the control of attention in securing the most favorable attitude of work, which would be involved in all types of mental work, (2) common principles in the assimilation and retention of the material, and (3) proficiency in reading. Problems involved in all of these elements would be: First, what are the specific processes common to all types of studying here referred to, and second, how may these various processes be facilitated?

Control of Attention. One of the chief, if not the chief, source of waste in studying and in fact in all mental work, is the reluctance in beginning an intellectual task. There seems to be in many individuals an almost insuperable inertia to overcome before work is, or can actually be, begun and continued without constant self-pushing. The common feeling is a dislike to begin work. "I don't like to study my history," or "I just hate to write this theme," "I don't see why he makes us do this," represent states of mind frequently found among the average pupils and to some extent even among the better pupils who often have severe struggles with such a tendency.

In papers on "Difficulties and Hindrances in Studying and How to Overcome Them" collected by the author from about one hundred university juniors and seniors, 56 mentioned lack of concentration, 26, dislike for or lack of interest in the subject, 23, getting started, 9, mind-wandering, 5, failure to organize material, and 4, day dreaming. These may all be classed as internal psychological difficulties centering around the problem of getting the mind to work at the task. Practically every student mentioned one or another or several of these four difficulties.

Besides attributing this situation to indolence or to stupidity, is there anything in the way of concrete suggestions and help that can be given to overcome this mountain of difficulty? I believe there are two general procedures which may be followed. One is to grit one's teeth and to "go to it"; that is, simply to force oneself by sheer voluntary effort to begin the task. The other is to put oneself into physical surroundings and into a frame of mind in which it will require a minimum, or at least a smaller amount, of voluntary effort. Strictly voluntary effort consumes a large amount of mental energy and, if it must be continued for a long time, is very wasteful of one's strength. The second is distinctly the more advisable plan to adopt. With the help of such a control of external conditions as is possible, the following means of directing one's energy may, therefore, be suggested:

(1) Put yourself into the proper physical or bodily attitude of

work. Sit up to your desk or table at which you customarily work. This in itself will help to start the mental machinery agoing and make it easier for the mental processes to operate.

(2) Work in surroundings in which there are absolutely no distractions as far as possible. Some persons can work under very distracting conditions, but these are exceptions, and if one has difficulty in beginning work, he should go alone into a separate room, shut the door, and sit facing away from the windows, and have nothing to look at or to attract his attention. A certain life insurance agent of one of the largest companies in America adopted the plan of selling to no one except by special appointment in his own office from which all possible distractions had been removed. There was nothing on the walls and nothing in the room but a desk, a telephone, and a couple of chairs. There was nothing on the desk except a life insurance policy, which was placed there at a certain time of the interview. The purpose was to secure conditions under which there were absolutely no distractions whatever, and the only thing to think about was the purchase of a life insurance policy. For a time there was a calendar of the company hanging above the desk. He found that many clients would remark, upon leaving the office, about the interesting dates designated on the calendar. There was nothing else to distract their attention and consequently these stood out in the minds of the clients, and, therefore, appeared interesting. He then removed the calendar to a rear wall so that even the dates might not distract. All these features were a part of his carefully prepared sales plan. This man was one of the most successful life insurance salesmen among all the agents of that company. In a certain month he had the record of selling the largest number of policies of all the salesmen of this large company—a record that was achieved after only eighteen months of experience in selling life insurance policies, immediately after graduation from college. It would, of course, be absurd to attribute his remarkable success to this one element, but it was nevertheless a very important part in a carefully prepared plan of salesmanship.

The removal of distractions, or what amounts to the same thing, the selection of a place for study where there are no distractions, is one of the most useful suggestions that anyone can adopt for developing concentration in work without a constant and exhausting tax upon the worker's voluntary efforts. In the course of time, it may be possible to work under even distracting circumstances, but probably no one, except the rare, absent-minded genius, can work as well among distractions of sights and sounds, and in the presence of other people, as under the complete absence of such stimuli. No one is in a position to appreciate the great effectiveness in intellectual work under complete absence of distractions until he has tried it. The average pupil wastes an inestimable amount of time by having to study in the presence of other members of the family who may be conversing or moving about, and every word or action or stimulus of any sort is bound to enter the mind and to divert the association processes to something else. Even though they are very minor, they require a few seconds, if not longer, to cause the thought process to return again to the subject-matter to be studied. Only those persons who have compared their own working efficiency under distracting conditions with their efficiency under ideal conditions can appreciate the enormous difference in the amount that can be accomplished.

(3) Begin work. Don't continue to think, "Oh, I just hate to do this," but instead go to your desk in your secluded room, sit down, take hold of book, paper, pencil, or whatever may be needed, and begin to write, or read, or figure. In short, if you have difficulty in overcoming inertia, just begin to go through the motions of work. This will automatically start the mental processes going relative to the work to be done, and before you realize it, you will be in the midst of the task, reading, thinking, and writing in an interested manner concerning the problems at hand. The external mechanical movements will act as stimuli for the inauguration of associative processes, and are likely to start mental activities without a great deal of voluntary effort.

A prominent story writer relates that he had difficulty in beginning his writing and in working out his plan necessary to finish up the details after the plot of the story had been conceived. This aspect of story writing is work and probably not a matter of inspiration; it involves close application and sometimes drudgery. He found that he was able to get into his writing by simply sitting down, taking a pencil and paper, and beginning to write whatever came to his mind, whether it was very pertinent to his story or not. Going through the motions started his thought processes agoing, and very shortly his associative and imaginative processes were almost automatically producing pertinent and excellent ideas.

In like manner, begin to study a lesson by taking the book, turning to the page, and simply looking at the print. Some voluntary effort, of course, must be exercised, if only to take hold of the book, but it is more economical to do so than continuously be thinking "How I hate to do it." This thought will automatically be driven out by the processes started by simply going through the motions of beginning work. The more voluntary effort and force one may be able to exercise in not thinking about dislike for the task and in beginning the motions of the work, probably the better and the sooner one is able to start, but this voluntary attention should normally pass very quickly into automatic attention and interest.

In the papers previously mentioned, the students stated that they overcame their difficulties of going at their tasks and keeping at them, besides "exercising will-power" which was mentioned most frequently, by "setting a certain hour to begin," by "doing work in a limited time," by "doing the work under pressure," by "dividing number of pages so that they could tell how many would have to be read every fifteen minutes," by "copying a sentence which helps to get the mind on the subject," by "starting directly for if I wait at all a million things would come up which were more interesting," by "having a definite schedule of study," by "planning the day," by "repeating with lips what is read," by "reading aloud," and by "studying in one place."

Common Elements in the Assimilation and Retention of the Material. At least five or six specific suggestions applicable to any kind of studying may be given.

(1) Take a problem solving attitude. Know definitely what you want to find. Ask questions and then look for the answers.

(2) Understand what you want to assimilate and retain permanently. To go through reading matter in a perfunctory manner will not cause retention of it except after long, wasteful, and frequent re-readings. A certain psychologist, in conducting experiments in memory with words and syllables, had dictated over and over a great many times certain series of materials, so that they had been completely memorized by several subjects, but he himself was unable to repeat the material from memory. The reason was that he himself had never paid strict attention to the memorizing, and had read them over and over again purely in a passive, inattentive manner.

Do not try to memorize ideas blindly; memorize understandingly. Some material in school must be memorized mechanically, but much more of it can be learned with a thorough conception of its meaning.

(3) Organize ideas with reference to certain larger ideas and principles. Organize your ideas and think out their relation to general principles. Grasp in as large units as possible and note the relation of details.

In the writer's *Experiments* ('17), page 190, is given an experiment in which two series of facts of apparently equal kind and difficulty are presented for memorizing. Each list is composed of five dates of history, five Greek words with their English meanings, and five numbers with their sums. In one list these ideas are arranged in miscellaneous order; in the other list they are grouped by subjects, the five historical dates are in one group, the five Greek words in the second group, and the five sums in the third group. The time required by ten subjects for memorizing the first set was an average of 14 minutes and 3 seconds; the time required for memorizing the second series, which was arranged in order, was on the average 9 minutes and 11 seconds. The comparison shows a very decided advantage in favor of learning the material in organized form.

(4) Recall at brief intervals the essential ideas of what you have read. Stop at the end of each paragraph or two, shut your book or your eyes, and recall the essential ideas you have read. Say to yourself "What did I read about?" Then try to answer the question. Note here what was said about forgetting in the last chapter. The chief value of examinations is the occasion and stimulus which they afford for recalling and organizing the material covered. In some respects the most valuable studying done by pupils is done in preparation for examinations. The value of the principle of recall in learning or memorizing has been thoroughly demonstrated by laboratory experiments.

Then each day or two, relate the recent material in a given subject to the earlier material in that subject. That is, review in your mind at short intervals, the larger essentials of all the material covered up to date. The principle of recall in this form is used far too little in studying. These suggestions would be applicable to every type of reading which has to be done rather carefully. It would, of course, not be advisable to do so in materials such as a novel in which the ideas in detail need not be retained.

(5) At the earliest possible moment and as frequently as possible, use the ideas that have been acquired, either by telling them

to some one else, or by thinking them over in your mind in connection with other related materials or situations. This will give them meaning in new ways and from new angles, and will help to fix them permanently by virtue of the principle of recall.

(6) In committing material to memory, learn by wholes rather than by parts. Poetry or prose can, as a rule, be memorized more quickly if the material is read through as a whole from beginning to end than if it is memorized in small sections of two or three lines; and what is more important, when this method is employed, the retention is more permanent. With many persons who are accustomed to memorizing by the part method, there is frequently no saving of time in the first learning partly because the whole method is new to them and partly because the learner often doubts the advisability of using the whole method.

There are three reasons why the whole method proves in the long run to be more economical: (1) Learning by parts establishes many useless and interfering connections. Thus in committing the first two lines of a poem the association is established between the last word of the second line and the first word of the first line. But this is not the order in which the lines are to be recalled. Rather the connection should be established between the last word of the second line and the first word of the third line as is done in the whole method. Consequently every portion memorized by itself forms at least one detrimental connection and in a long selection a very considerable number of such associations are formed. These derailing paths probably account for the fact that pupils in reciting a poem become stalled usually between the portions learned piecemeal. (2) Reading the material over as a whole gives a view of the entire selection and will serve to give meaning and correlation of the parts in the whole. It will help to organize the ideas as a whole. (3) Learning by parts is apt to produce great unevenness among the various portions of the material in the degree of perfection of the memorizing. Some parts, especially the earlier ones, will be repeated needlessly a great many times and result in much greater over-learning of those parts than of other parts. One point of caution in using the whole method should, however, be noted. When the learner reads over the entire selection to be memorized he does not make much visible progress until, after a sufficient number of repetitions, he is able to reproduce most of the material. This situation is likely to be discouraging, particularly to children. Perhaps the most effective manner of employing the whole method is to learn the material in relatively large sections instead of as a complete whole, particularly if the selection is very long.

Improvement in Reading Ability. The average child, as well as the average adult, reads far too slowly, and in fact, far more slowly than he is capable of reading. About one-fourth of university students read less rapidly than the average 8th grade pupil does, and about one-fourth of 8th grade pupils read less rapidly than the average 5th grade pupil. Experiments indicate that by a moderate amount of definite practice, with conscious effort to improve, the speed of reading may be increased from 50% to 100% without loss in the comprehension of the ideas read. The moral would be: Force yourself to read more rapidly, which will be accompanied by greater concentration of attention and in the course of time this more rapid reading will become habitual, so that the comprehension will be just as complete as at the slower rate of reading. Consult the latter part of the chapter on "Reading" for a more detailed discussion of these points.

Concrete Rules for Studying. Whipple has presented a series of thirty-eight rules which ought to prove valuable for increasing effectiveness in studying. Some of these rules involve points that have been previously presented in this chapter. Their specific character makes them commendable for the student's consideration and observance. They are as follows:

SUMMARY OF RULES. After Whipple ('16)

1. Keep yourself in good physical condition.
2. Attend to, remove or treat physical defects that often handicap mental activity, such as defective eyesight, defective hearing, defective teeth, adenoids, obstructed nasal breathing.
3. See that external conditions of work (light, temperature, humidity, clothing, chair, desk, etc.) are favorable to study.
4. Form a place-study habit.
5. Form a time-study habit.
6. When possible, prepare the advance assignment in a given subject directly after the day's recitation in it.
7. Begin work promptly.
8. Take on the attitude of attention.
9. Work intensely while you work: Concentrate.
10. But don't let intense application become fluster or worry.
11. Do your work with the intent to learn and to remember.
12. Seek a motive or, better, several motives.

13. Get rid of the idea that you are working for the teacher.
14. Don't apply for help until you have to.
15. Have a clear notion of the aim.
16. Before beginning the advance work, review rapidly the previous lesson.
17. Make a rapid preliminary survey of the assigned material.
18. Find out by trial whether you succeed better by beginning with the hardest or with the easiest task when you are confronted with several tasks of unequal difficulty.
19. In general, use in your studying the form of activity that will later be demanded when the material is used.
20. Give most time and attention to the weak points in your knowledge or technique.
21. Carry the learning of all important items beyond the point necessary for immediate recall.
22. You must daily pass judgment as to the degree of importance of items that are brought before you, and lay special stress on the permanent fixing of those items that are vital and fundamental.
23. When a given bit of information is clearly of subordinate importance and useful only for the time being, you are warranted in giving to it only sufficient attention to hold it over the time in question.
24. Make the duration of your periods of study long enough to utilize "warming-up" but not so long as to suffer weariness or fatigue.
25. When drill or repetition is necessary, distribute over more than one period the time given to a specified learning.
26. When you interrupt work, not only stop at a natural break, but also leave a cue for its quick resumption.
27. After intensive application, especially to new material, pause for a time and let your mind be fallow before taking up anything else.
28. Use various devices to compel yourself to think over your work.
29. Form the habit of working out your own concrete examples of all general rules and principles.
30. Form the habit of mentally reviewing every paragraph as soon as you have read it.
31. Don't hesitate to mark up your own books to make the essential ideas stand out visibly.
32. Whenever your desire is to master material that is at all extensive and complex, make an outline of it. If you also wish to retain this material, commit your outline to memory.
33. In all your work apply your knowledge as much as possible and as soon as possible.
34. Do not hesitate to commit to memory verbatim such materials as definitions of technical terms, formulas, dates and outlines, always provided, of course, that you also understand them.
35. When the material to be learned by heart presents no obvious

rational associations, it is perfectly legitimate to invent some artificial scheme for learning and recalling it.

36. In committing to memory a poem, declamation or oration, do not break it up into parts but learn it as a whole.

37. In committing to memory, it is better to read aloud than to read silently and better to read rapidly than slowly.

38. If your work includes attendance at lectures, take a moderate amount of notes during the lectures, using a system of abbreviations, and rewrite these notes daily, amplified into a reasonably compendious outline, organized as suggested in Rule 32.

Supervised Study. Teachers have come to recognize in recent years the waste of time and the blind direction of energy, or possibly lack of energy, in so much of the studying done by pupils that a widespread movement has gotten under way for the supervision of studying. The plans for supervising studying are carried out in so many different ways that hardly any one plan can be designated as typical. The results accruing from the general efforts in this direction have been in most cases beneficial. Continued experimentation during the next few years with various plans of supervised study will lead to a more general agreement as to the most effective manner of administering it.

In a recent inquiry of supervised study in schools on the Pacific coast, Proctor ('17) found that forty-two high schools employed it in one form or another. Of these forty-two schools, thirty-one reported the use of a lengthened period distributed as follows:

(a) 60′ period, divided 30–30, No. of cases......................	3	
60′ period, divided 35–25, No. of cases......................	1	
60′ period, divided 40–20, No. of cases......................	15	
60′ period, divided 45–15, No. of cases......................	1	
63′ period, divided 33–30, No. of cases......................	1	
	—	21
(b) 70′ period, divided 40–30, No. of cases......................	4	
70′ period, divided 35–35, No. of cases......................	2	
	—	6
(c) 80′ period, divided 40–40, No. of cases......................	1	1
(d) 85′ period, divided 45–40, No. of cases......................	2	2
(e) 90′ period, divided 45–45, No. of cases......................	1	1
		—
Total..		31

Regarding the effects of supervised study, Proctor reports that:

"Twenty-six of the 31 principals employing the lengthened period said that study habits had been improved; one could discover no apparent effect; two said that only the slow students had been helped, the brighter ones were not; and two had no data on which to base their opinions.

"Wherever the plan had been in use long enough to make possible the compiling of statistics as to the effect of supervised study on scholarship, there was practically unanimous agreement that the number of failures had been reduced and the standards of scholarship had been raised. The high school at Snokomish, Washington, reports that the average percentage of failures in elementary algebra for the two years prior to the adoption of supervised study was 28%. But for the two-year period following the adoption of supervised study the failures in the same subject were reduced to 17%. Hoquiam, Washington, reports that the average marks of the students range 10% higher and that the number of honor pupils has been doubled since supervised study was introduced. The principal of the Arcata high school, California, reports that the average mark of the freshman class has been raised from 78% to 82½% during the first year of supervised study. Santa Cruz, California, comparing the year 1914–15, the last under the old plan, with the year 1916–17, the second year under supervised study, finds that the increase in the total number of high marks has been 157%; the decrease in low failures, 188%. Reno, Nevada, reports a decrease of 45% in the number of failures, and an increase of 24% in the number of students making excellent marks."

J. Stanley Brown, principal of the high school at Joliet, Illinois, reports, as quoted by Hall-Quest ('17), a decided reduction in the percentage of failures after the introduction in the high school of supervised study, as indicated in the following table:

TABLE 46. After Brown and Hall-Quest ('17, p. 386). Supervision of study apparently was begun in 1912 although I have not been able to find a definite statement by Hall-Quest to that effect.

Table of percentage of failures

Subject	1911	1912	1913	1914
Algebra	24	22	15	12
Arithmetic	26	20	12	13
Geometry	29	19	17	16
German	21	20	13	14
Latin	22	19	16	13
French	10	9	8	9
Physiography	12	10	8	9

Breslich ('12) made an experiment to determine the effect of directed study by dividing an algebra class into two sections, one of which was conducted in the usual manner of recitation work and home study, and the other was conducted by confining all of the work to the recitation period of 45 minutes. This time was devoted partly to study under the supervision of the teacher and partly to recitation work. The two sections were made up of pupils of approximately equal ability as indicated by the marks for the preceding semester's work, which averaged 81.4 for those who constituted the home study group, and 79.4 for those who constituted the supervised study class. The home study class devoted approximately two hours to each lesson including the 45 minutes for the recitation period. The experiment was conducted for a period of fourteen lessons. At the end of that time, the same examination over the work that had been covered was given to

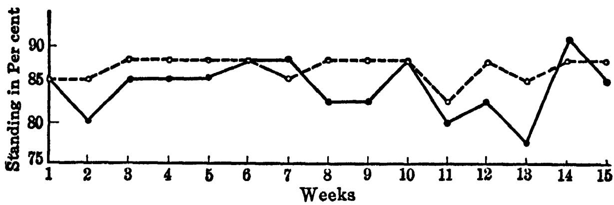

FIG. 54.—The broken line represents the supervised group. The continuous line represents the unsupervised group. After Minnick ('13).

both sections. The supervised study class made an average of 65.5, and the home study group made an average of 62.8. Hence the supervised study class obtained as good results as the home study class, or slightly better, in spite of the fact that the former spent only about two-fifths as much time upon the work.

Minnick ('13) of Bloomington, Indiana, divided a group of thirty-six pupils in plane geometry into two divisions, providing supervision of the study to one division and none to the other division for a period of fifteen weeks. The results show an advantage in scholarship for the supervised group as represented in the curves of Figure 54.

Hall-Quest quotes similar results of improvement in scholarship as indicated either in the reduction of the percentage of failures or in the higher scholastic marks as reported by Loveland at Pottstown, Pennsylvania, by Rickard, at Oakland City, Indiana, and from the high school at Pueblo, Colorado.

CHAPTER XIII

TRANSFERENCE OF TRAINING IN SPECIFIC MENTAL FUNCTIONS

Problems. To what extent does training in one mental function or set of functions modify the operation of other mental functions? To what extent will training in mathematical reasoning modify reasoning ability about political events or bargains, or vice versa? To what extent will training in remembering faces and names modify the remembering of the prices of goods or words in a language, or vice versa? To what extent will improvement in proficiency in such mental capacities as are involved in school studies modify proficiency in any other specific or general activities or interests in life? These problems open up all the ramifications of the traditional controversy concerning mental discipline, the educational value of school subjects, and their related discussions. The fundamental problem, however, is not, does training transfer? The task is more complex and suggests rather the following three fundamental problems: (1) To what extent does training transfer, (2) To how closely or how distantly related functions does training transfer, and (3) How does the transfer take place? Thus we see that transference of training is one of the three or four most important perennial problems in the entire field of education. So many problems in the administration of schools, in the construction of courses of study, in the emphasis upon various aspects of school subjects, in fact, the ultimate values of education as a whole, depend primarily upon our attitude toward the problems of mental discipline and transference of training.

In order to think about these matters clearly, it is necessary to distinguish between two quite different aspects of the discussion, namely, (1) The pure disciplinary or training value in the improvement of a mental function irrespective of the material through which it is trained or which is acquired in the training; (2) the intrinsic value of the information or material acquired in the process of the training irrespective of the training referred to in (1). The one is acquisition of training; the other is acquisition of content. The difference between these two aspects may be illustrated thus: The learning of shorthand will furnish practice of certain types of memory and associative processes; it will also supply the individual with certain symbols for recording ideas. The former would be the pure training value of the mental functions, the latter would be the informational or instrumental value of shorthand. If one should never expect to use the symbols of shorthand for recording ideas, to what extent would the practice in memory and associative functions modify memory and associative processes in other reactions in life? How much value, accordingly, may we attach to the practice of these mental processes? In thinking about these problems, we must distinguish sharply between the content or informational aspect of a given type of learning and the pure improvement value in mental functions to be derived from the learning. Viewed from the standpoint of the school, the situation presents two problems: (1) To what extent does training of the mental capacities involved in a given school subject carry over and produce efficiency in other subjects or in other activities in life, and (2) are certain school subjects more capable of improving the mental functions generally and of carrying the improvement over to other responses of behavior?

The Effect of Improvement in Specific Mental Functions upon other Mental Functions. The influence of improvement in one function upon others may be one of help, hindrance, or indifference. Which it is and how much, can be determined only by recourse to facts. Until twenty-five years ago the problem was discussed wholly as a matter of opinion. During the last twenty-five years, a considerable number of researches have been made on many aspects of the problem so that the controversy may be dealt with in a more definite and factual manner than was formerly the case.

The experimental technique of research in the field of transference of training has been practically the same in all investigations, and has consisted (1) of testing the strength of a variety of mental capacities, (2) of training one or more capacities for a specified period of time, and (3) by finally testing again the same capacities tested before the training in order to determine what changes may have been produced in them as a result of the intervening training. The tests referred to under (1) and (3) are conveniently called "end tests" or the "test series" and the work under (2) is usually called the "training series." This plan has been followed in the large majority of transfer experiments. A different plan, however, is possible and has been employed in a few studies. This consists of giving training to a group of persons in some particular function and then giving them practice in another function. Their progress in this second function is then compared with that of other individuals who have not had training in the first function.

a. James' Experiment on Memory. The first experimental investigation was made by James and published in 1890 in his *Principles of Psychology.* This experiment is of interest and importance chiefly because of its historical significance in opening the problem by an experimental approach. James attempted to determine the effect of training in learning one kind of poetry upon memorizing other kinds of poetry. He first made the experiment upon himself by memorizing in the course of eight days 158 lines of Victor Hugo's *Satyr.* This required a total of 131 $^{5}/_{6}$ minutes. He then spent some twenty minutes a day for 38 days in learning the first book of Milton's *Paradise Lost.* At the end of this time he again memorized 158 lines from Victor Hugo and found that it took 151½ minutes. This loss in time was surprising and James explained it by saying that he was fagged out by other work at the time of the second test on Victor Hugo and that he was not really in fit condition for such an experiment. He then repeated the experiment with four students in a similar manner by using different poetry. The results of these early experiments are given in the following table:

TABLE 47. After James ('90, I, p. 667)

Individual	Test Before Training	Training	Test After Training
1	158 lines of Victor Hugo during 8 days, 131.8 minutes	1st Book Paradise Lost 38 days	158 lines of Victor Hugo during 8 days, 151.5 minutes
2	128 lines of In Memoriam during 8 days, 14.8 min. daily average	416 lines Schiller's translation of the Æneid during 26 days	128 lines of In Memoriam during 8 days, 14.6 daily average
3	? of Virgil during 16 days, 13.4 min. daily average	? of Scott	? of Virgil during 16 days, 12.3 min. daily average
4	150 lines of ? during 15 days, 3.7 min. daily average	450 lines of ?	150 lines of ? during 15 days, 3.0 min. daily average
5	? lines of Idylls of the King during 6 days, 14.6 min. daily average	? of Paradise Lost	? lines of Idylls of the King during 6 days, 14.9 min. daily average

This experiment in its essential details, was repeated by Peterson ('12) with two subjects, one of whom showed gain and the other loss.

b. Reaction Time. The next series of experiments was undertaken by Gilbert and Fracker, who attempted to determine the amount of transference of training from one type of reaction to other types of reaction. Three subjects were tested first in simple reaction to sound, to electric stimuli, to touch, to visual stimuli, and likewise in complex reaction to stimuli involving discrimination and choice. The training series consisted of simple and complex reaction to sound only and continued for twelve days. The results obtained in this experiment are given in the following table which shows the percentages of gain made in each of the end tests:

TABLE 48

The spread of improvement in reacting to various sensory stimuli. After Gilbert and Fracker ('97)

Individual	The Percentages of Time Gained by Practice							
		Simple Reaction			Reaction with Discrimination and Choice			
	To Sound	To Electric Shock	To Touch	To Color	Sounds	Electric Shocks	Touch	Blue and Red
J. A. C.	12	−2	17	3	53	35	9	14
G. C. F.	23	21	10	45	47	60	38	34
J. C. P.	13	16	6	11	14	24	4	19
Averages	16	17	11	20	50 [1]	40	17	22

[1] Average of J. A. C. and G. C. F. only.

J. C. P. was practiced only in reaction time, while the other two were practiced in both reaction and reaction with discrimination and choice. All figures of the above table represent per cent of gain by practice.

Each of the forms of reaction shows on the whole a distinct gain in the second end tests. How much of this gain is actually due to the training series cannot be definitely determined. Many of the earlier investigators did not make control tests, that is, they did not repeat the end tests on another group of subjects who did not take the practice series but who took only the end tests separated by an interval equal to that consumed by the practice series. It is obvious that a certain portion of the gain in the end tests is due to the fact that when the second end tests are made, some advantage is derived from the familiarity or practice in having done the end tests once before; consequently, the actual amount of improvement in a practice experiment can be determined only by subtracting the amount of gain made by a control group which has not done the practice series in order to obtain the residual amount of improvement actually transferred from the training series.

Another important item frequently omitted in the early investigations is a statement of the actual amount of progress made in the practice series itself. This element is significant because it is possible thereby only to determine the amount of gain made in the end tests as compared with the improvement in the training series itself in order that some definite conception may be formed of the amount of gain made in the practice series which is transferred to the end tests. Thus in the reaction experiments of Gilbert and Fracker, the gain in the practice series is shown in the first and fifth columns. It will be noticed that the average gains in the end tests in simple reaction to electric shocks, to touch, and to color was about as great as in the training series itself, that is, in simple reaction to sound. It was 17%, 11%, and 20%, or on the average 16% in the former, as compared with 16% in the latter. In case of the complex reactions, the average gains in the reactions to electric shocks, to touch and to color were 40%, 17%, and 22%, or on the average 26%, as compared with a gain of 50% in the practice series. On the face of it, 100% of the practice effect in simple reaction to sound was transferred to the other forms of simple reaction, while 52% of the practice effect in complex reaction to sound was carried over to the other types of complex reactions. Actually the amounts of transfer effects are probably considerably less; how much we do not know since Gilbert and Fracker made no control tests.

c. Perception and Discrimination. Thorndike and Woodworth ('01) made an investigation to determine the transference of practice in estimating areas, lengths of lines, and weights to estimating areas, lines, and weights of different sizes. They also measured the effect of practice in perceiving words containing certain letters upon the accuracy and quickness of perceiving other words containing different letters. The results of this experiment are summarized in the following manner by Thorndike:

"Individuals practiced estimating the areas of rectangles from 10 to 100 sq. cm. in size until a very marked improvement was attained. The improvement in accuracy for areas of the same size but of different shape due to this training was only 44% as great as that for areas of the same shape and size. For areas of the same shape, but from 140–300 sq. cm. in size, the improvement was 30% as great. For areas of different shape and from 140–400 sq. cm. in size, the improvement was 52% as great.

"Training in estimating weights of from 40–120 grams resulted in only 39% as much improvement in estimating weights from 120 to 1800 grams. Training in estimating lines from .5 to 1.5 inches long (resulting in a reduction of error to 25% of the initial amount) resulted in no improvement in the estimation of lines 6–12 inches long.

"Training in perceiving words containing 'e' and 's' gave a certain amount of improvement in speed and accuracy in that special ability. In the ability to perceive words containing 'i' and 't,' 's' and 'p,' 'c' and 'a,' 'e' and 'r,' 'a' and 'n,' 'l' and 'o', mispelled words and A's, there was an improvement in speed of only 39% as much as in the ability specially trained, and in accuracy of only 25% as much. Training in perceiving English verbs gave a reduction in time of nearly 21% and in omissions of 70%. The ability to perceive other parts of speech showed a reduction in time of 3%, but an increase in omissions of over 100%."

The experiments in marking out words and in estimating weights were repeated with two persons in substantially the same manner by Coover. ('16.)

"Two reagents were trained for 11 days in marking out words containing e and s in selected columns of the 'Outlook' Magazine. Each reagent looked over 12,000 words in each day's practice.

"Tests were taken before and after training, in marking out

"(1) Words in 'Outlook' columns containing e-s, i-t, s-p, c-a, e-r.

"(2) Words on manuscript pages containing a-n, l-o, e-r.

"(3) Common nouns in 'Outlook' columns.

"(4) Words in 'Outlook' columns containing e-s."

Coover's results showed a gain of 44% in the training series and of 33% in the end tests, or 75% as much as in the training series. This is a larger transfer effect than that of Thorndike and Woodworth whose results, however, were based on five persons and showed a gain of 37.7% in the training series and of 17% in the end tests, or 48% as much as in the training series.

Coover's experiment in estimating weights was carried out by training two persons with a set of seventeen blocks ranging from 40 to 120 grams. Each person made 1,700 judgments. The persons were tested, before and after the training, in estimating ten common objects averaging 67.5 grams in weight but falling within the limits of 40 and 120 grams, and in estimating ten common objects averaging 552.7 grams but all exceeding 120 grams.

The experiment yielded a gain in accuracy of estimating weights of 23% in the training series and of 29% in the end tests with the set of ten smaller objects but a loss of 100% with the larger objects. This loss was due to the very large loss of one subject which far outweighed the gain of the other subject. The gain in the estimation of the smaller weights was greater than in the training series itself. Thorndike and Woodworth's experiments showed a gain of 45% in the training series and of 38% in the end tests with the smaller weights and of 16% with the larger weights.

Kline had nine persons practice for fourteen days from 30 to 45 minutes daily in canceling e's and t's on pages of prose. Before and after the practice he tested them in canceling nouns, verbs, prepositions, pronouns, and adverbs. Eight other persons were tested in like manner without doing the practice series. Kline found that the practiced group did not gain as much as the unpracticed group. This he explains by the introspective statements of his subjects that "there was a tendency to cross out words containing e's and t's rather than the required part of speech." The detailed results follow:

TABLE 49

The spread of improvement in marking letters. After Kline ('09, p. 10)

	Nouns			Verbs			Prepositions			Pronouns			Adverbs		
	Correctly Marked	Wrong Words	Omitted	Correctly Marked	Wrong Words	Omitted	Correctly Marked	Wrong Words	Omitted	Correctly Marked	Wrong Words	Omitted	Correctly Marked	Wrong Words	Omitted
Practiced Group															
After practice	34.0	1 6	12.6	11.4	5.0	6.0	28.0	0.5	7.2	8.5	2 3	6.3	3.5	0.6	6 3
Before practice	28.6	4 6	17.3	9.8	6.5	4 0	25.9	3.0	8.2	6.0	4 4	5 0	6 6	1.7	9 3
Differences...	7.4	3.0	4.7	1.6	1.5	−2.0 [1]	2 1	2.5	1.0	2.5	2 1 [1]	−1.3 [1]	3.1	1.1	3.0
Unpracticed Group															
Second period	30.4	1.4	10.3	11.3	6.0	7.0	26.6	1.7	9.3	5.0	0.6	4.0	5.5	0.7	7.0
First period..	23.5	5.1	17.0	8.7	7.0	5.0	16.6	2.6	10.5	4.6	0.3	13.7	4.4	2.0	13.0
Differences...	6.9	3.7	6.7	2.6	1.0	−2.0 [1]	10 0	0.9	1.2	0.4	−0.3 [1]	9.7	1.1	1.3	6.0

[1] — sign indicates loss at second period.

Bennett ('07) tested a group of sixteen pupils in discriminating between shades of red, yellow-green, and orange, and differences in the pitch of tones before and after training twice a week for five months in discriminating shades of blue. The accuracy in the four end tests showed the following gains:

	1 Red	2 Yellow-green	3 Orange	4 Tones
Boys........	79%	60%	65%	28%
Girls.........	84%	57%	56%	23%

Coover and Angell tested four adults in discriminating intensities of brightness before and after training in discriminating intensities of sound consisting of seventeen series of forty judgments each. The end tests without the intervening training were also given to three other subjects. The four trained persons rose from 56.9% of right judgments before the training, to 66.0% after the training, while the untrained persons dropped from 65.5% right judgments to 61.7%.

d. Sensori-motor Association. Bair ('02) attempted to measure spread of practice, not by testing certain capacities before and after training in some other capacity, but by training the subjects in a certain function and then determining the effect of this training upon the progress in the subsequent training of other functions. His experiments are described thus:

"(1) Six keys of a typewriter are labeled with six symbols (letters or figures). Fifty-five of these letters or figures, in chance order, are now shown one by one, and the subject on seeing one taps the corresponding key. The time taken to tap out the series is recorded. Six different symbols are then used with a new series composed of them, and the subject's time record is taken as before. This is continued until twenty different sets of symbols have been used. Although the symbols have been changed each time, there is a steady improvement, ranging for the four subjects in the following decrease in time: 62 to 52, 95 to 85, 71.5 to 58, 65 to 56. The major part of this gain could not have been due to merely getting used to the machine or to the general features of the experiments, for the fourth subject was already used to these and still gained about nine-tenths as much as the other three.

"(2) The other experiment consisted in taking daily records for twenty days, by means of a stop-watch, of the time required to repeat the alphabet from memory. Each day's experiment was as follows: First, the alphabet was repeated as rapidly as possible forward; second, the letter n was interpolated between each of the letters; third, the alphabet was repeated backward interpolating n between each two of the letters. At the end of twenty practices in each order the subject repeated the alphabet first forward interpolating instead of n the letter x and repeating three times; secondly, interpolating r and repeating three times; then lastly, repeating backward and in like manner interpolating x and r and repeating three times. There was improvement in the test series, the effect of the twenty days' training with the training series being to put the abilities in the test series as far ahead as three days of the direct training would have done."

Scholckow and Judd investigated the effect of knowledge of the principle of refraction upon learning to hit a target under water.

"One group of boys was given a full theoretical explanation of refraction. The other group of boys was left to work out experience without theoretical training. These two groups began practice with the target under twelve inches of water. It is a very striking fact that in the first series of trials the boys who knew the theory of refraction and those who did not, gave about the same results. That is, theory seemed to be of no value in the first tests. All the boys had to learn how to use the dart, and theory proved to be no substitute for practice. At this point the conditions were changed. The twelve inches of water were reduced to four. The differences between the two groups of boys now came out very strikingly. The boys without theory were very much confused. The practice gained with twelve inches of water did not help them with four inches. Their errors were large and persistent. On the other hand, the boys who had the theory, fitted themselves to four inches very rapidly." (Judd, '08, p. 37.)

Webb ('17) used the plan of determining the effect of acquired skill upon the acquisition of other skills. He employed 54 rats and 21 humans in learning mazes in various orders. He measured the results in terms of the number of trials required, the number of errors made, and the amount of time needed to learn the mazes. The following table gives the savings in learning a second maze as compared with the learning of the first one:

TABLE 50. After Webb

Average percentage of saving in transfer

Rats				Humans			
Mazes	Trials	Errors	Time	Mazes	Trials	Errors	Time
A—B...	77.08	85.81	83.77	A—D...	51.98	94.58	88.73
A—D...	69.02	79.71	90.42	A—B...	67.86	86.64	67.18
A—E...	19.91	54.63	63.40	A—C...	19.74	20.20	29.18
A—F...	63.01	42.78	59.44				
A—C...	57.85	46.10	34.94				

Webb concluded that "the learning of one maze has a beneficial effect in the mastery of a subsequent maze situation" (page 50) and that "the degree of transfer is dependent in part upon the degree of similarity of two maze patterns" (page 53). Webb further attempted to ascertain whether a new habit has a retroactive effect upon habits previously formed. He had his subjects learn one maze, then a second one, and then return to the first one. His findings were inconclusive.

Coover ('16) reports an unpublished investigation by Carrie W. Liddle designed to measure the effect of practice in discriminating and sorting cards bearing colors or geometric signs upon discriminating and sorting cards with different colors or signs.

"Each set of 102 cards contained six colors, or six designs, was shuffled so that no color or device repeated itself, and was sorted into six compartments. The first six cards of the pack determined the order of colors in the compartments according to which the rest of the pack was to be sorted. Nine reagents took part and the experiment continued two semesters. There was transference of practice-effect from one set of colors to the other set of colors, and to the geometric forms; and from one set of geometric forms to the other and to the colors. Increased powers of discrimination and attention were thought to be the causes of transference."

Bergstrom ('94) had found previously that training in sorting cards by one method interfered with sorting them by a different method. The same situation is shown by the card-sorting experiment in the author's *Experiments*, Chapter XV.

Coover and Angell ('07) attempted to ascertain the effect of practice in card sorting upon typewriter-reactions. They trained four persons in card sorting on 15 days scattered through a period of 40 days. During that time the subjects sorted 4,200, 3,800, 5,200, and 4,000 cards respectively. Before and after this training they were given practice in typewriter-reactions. Three other persons, as a control group, were given practice in typewriting at two periods separated by an interval of 45 days. The results are interpreted by the authors as indicating transfer, but it is doubtful whether there is any transfer and, if there is, how much. The practiced group reduced their time for the first 100 typewriter reactions, before the training in card sorting, from 84.4 seconds, with an average of 2.3 errors, to 62.3 seconds, with 6.3 errors, for the last 100 reactions after the training in card sorting. The unpracticed group reduced their time from 106.3 seconds, with 3.3 errors, to 80.6 seconds, with 2.3 errors. The trained group reduced its time by 26% but increased in errors, while the untrained group reduced its time by 25% but decreased in errors. There is obviously no appreciable transfer.

c. Memory. More extensive researches have been made in the field of memory than in any other single aspect of the problem of transference of practice. One of the most elaborate investigations was made by Ebert and Meumann. They measured the amount of transfer from memorizing a series of nonsense syllables to various other types of memory, such as immediate memory for numbers, letters, words, permanent memory of prose, poetry, etc. The end tests were made at three different times, before the beginning of practice series, about the middle, and at the close of the practice series. The results showed very considerable gains in these other types of memorizing. The difficulty in interpreting their results, however, is the fact that they did not make the cross section tests with a control group according to which a deduction could be made for the gain in the end tests themselves. Dearborn repeated the end tests on a group of subjects to ascertain the amount of allowance to be made. His results together with those of Ebert and Meumann are shown in the following table. Dearborn found a very considerable amount of gain in these end tests. His comments are as follows:

"The results indicate that a considerable part of the improvement found must be attributed to direct practice in the test series, and not to any 'spread' of improvement from the practice series proper. There is further, at times, lack of correlation between the amount of improvement made in the practice and that made in the test series; occasionally a larger percentage of gain is made in the latter than in the practice itself. This again indicates the presence of direct practice in the test series.

"Some at least of the remaining general improvement found is to be explained simply in terms of orientation, attention, and changes in the technique of learning.

"These results seem to render unnecessary the hypothesis proposed by Ebert and Meumann to account for the large extent of the general influence of special practice, which their experiments seem to indicate."

Three subjects of Ebert and Meumann were trained in learning 64 sets of 12-syllable series; they gained 70%. Three others were trained with 48 sets. They gained 50%.

TABLE 51. After Ebert and Meumann ('05) and Dearborn

End Tests	Ebert and Meumann's Gains of		Dearborn's 3rd Cross Section Test over the First[1]	Difference
	2nd Cross Section over the First	3rd Cross Section over the First		
Memory Span:				
Numbers	24%	60%	12%	48%
Letters	32	35	29	6
Syllables	22	43	17	26
Memorizing:				
10-syllable series	67	81	41	40
12- " "	61	76		
14- " "	64	80		
16- " "	48	70		
Geom. Forms (easy)	24	72		
" " (hard)	36	73		
German-Italian Vocab. (30 pairs)	—12	34	52	
German-Italian Vocab. (40 pairs)	—29	22		
Poetry 16 lines	6	11	14	—3
Prose 20 lines	42	72	58	14
				22%

[1] Quoted by permission from an unpublished table prepared by Dearborn.

Thus the high percentages of Ebert and Meumann are reduced to an average residual transfer of 22%.

Fracker has reported a rather extensive series of investigations on transfer in memory in which the training series consisted of memorizing various combinations of four degrees of loudness in a sound. These four loudnesses were presented in the various possible combinations and the responses of the subjects consisted in indicating the proper order in which the sounds had been received. The end tests consisted in determining the memory capacities for various combinations of four shades of gray, 9 tones, 8 shades of gray, 4 tones, geometrical figures, 9 sets of numbers, arm movements, and poetry. The results are summarized in the following table which also indicates the amount of deduction to be made due to the improvement in the end tests made by the control group.

TABLE 52

Transference of training in memory. After Fracker ('08)

The improvement made in Training Series by 8 subjects was 21%.

End Tests:					
Similar to Training Series:					
Four Grays,	8 trained subjects,	36%	4 untrained,	4%	32%
Nine Tones,	8 " "	22	4 "	11	11%
Nine Grays,	8 " "	19	4 "	10	9%
Four Tones,	8 " "	10	4 "	−2	12%
		22%		6%	16%
End Tests:					
Unlike Training Series:					
Geometrical Figures,	8 trained subjects,	13%	4 untrained,	8	5%
Nine Numbers,	8 " "	4	4 "	0	4%
Movement,	8 " "	0	4 "	−1	1%
Poetry,	8 " "	7	4 "	2	5%
		6%		3%	3%

An interesting result emphasized by these data is the fact that the transfer to the types of memory similar to that involved in the training series is considerably greater than the transfer to the memory functions unlike the training series. The average residual gain in the four similar memory processes is 16%, whereas in the four unlike memory processes it was only 3%.

Sleight made a careful and extensive investigation on transference of training in one sort of memory to other sorts of memory. He believed that previous researches had not used enough subjects to be statistically reliable. He therefore carried out his first research with 84 pupils from three girls' schools, averaging 12 years and 8 months old. Ten cross sectional tests were made before, in the middle, and after the training series, as follows: (1) Remembering and reproducing the location of points in circles, (2) two series of six dates each and their corresponding events, (3) series of eight syllables, (4) a stanza of from eight to twelve lines of poetry, (5) learning a passage of prose, (6) reproducing the content of a passage of prose, (7) remembering locations on a map, (8) remembering dictated sentences, (9) memory span for letters, (10) remembering names.

The pupils were divided into four groups of approximately equal ability as determined by the ten tests before the training series. One group was then trained in learning poetry; another in learning tables of multiplication, denominations, squares, fractions, etc.;

a third in reproducing the thought content of prose selections of scientific, geographical and historical material; and the fourth group had no special practice. The training period lasted four days a week for six weeks, practicing 30 minutes each day.

The chief results are presented in Table 53. I have computed the percentage of gain made by each group in Section III, that is, the end tests made at the close of the training series, over Section I, the tests made before the training series. These percentages are given in the last column. Sleight has not made such a percentage comparison, but has used a different, and possibly fairer, plan of computing the data. I have, however, made this computation in terms of percentages as these will be more intelligible to the reader unfamiliar with statistical methods. The average percentages at the bottom of the table show only slight gains on the part of the trained groups, 2, 3, and 4, over the untrained group. The average gain of group 2, trained in poetry, over group 1, untrained, was 3.3%; of group 3, trained in arithmetical tables, over group 1 was 2.6%; and of group 4, trained in prose, over group 1 was 4.0%. The amounts of transfer are very small. Sleight failed to indicate the improvement in the training series themselves so that it is impossible to compare the transferred amount with it.

TABLE 53

The numbers in the following table are the average scores made by each group in each test. Group 1 had no special practice, Group 2 was practiced in learning poetry, Group 3 in learning tables, and Group 4 in learning prose substance.

The column under Section I gives the scores before the training, under Section II about the middle of the training, under Section III after the training. After Sleight. ('11, p. 413.)

	Section I Early Test	Section II Middle Test	Section III Final Test	Percentage Gain of III over I
Points.....Group 1	73.9	86.2	86.5	17
" 2	66.8	80.2	84.5	25
" 3	66.5	77.2	90.3	26
" 4	58.5	69.8	76.5	31
Dates .....Group 1	14.4	15.3	18.1	26
" 2	14.7	16.8	20.4	38
" 3	18.9	21.9	21.3	13
" 4	17.7	17.1	20.1	14

TABLE 53—*Continued*

	Section I Early Test	Section II Middle Test	Section III Final Test	Percentage Gain of III over I
Nons. Sylls. Group 1	20.7	20.7	22.8	10
" 2	19.8	24.9	27.3	33
" 3	19.2	24.9	28.2	47
" 4	21.9	21.0	24.6	12
Poetry.....Group 1	58.5	62.4	63.8	9
" 2	56.5	59.4	57.9	3
" 3	60.3	60.9	64.4	.7
" 4	59.4	63.4	74.7	25
Prose (literal) Group 1	109.8	117.4	118.6	8
" 2	101.9	107.3	107.5	7
" 3	108.1	113.0	115.6	7
" 4	104.6	113.7	118.3	14
Prose Subs. Group 1	27 5	28.8	30.5	11
" 2	23 5	24.8	24.7	5
" 3	23.5	27.1	27.1	15
" 4	22.8	28.8	28.3	20
Map Test..Group 1	63.9	65.9	72.4	13
" 2	65.9	65.1	81.9	25
" 3	65.9	64.0	74.5	13
" 4	68.3	66.8	78.7	16
Dictation ..Group 1	134.1	135.9	139.0	4
" 2	129.6	130.9	130.0	0
" 3	129.3	130.3	132.8	3
" 4	129.8	133.6	134.7	4
Letters ....Group 1	76.1	78.9	80.2	5
" 2	79.2	81.7	82.6	4
" 3	76.5	78.4	80.8	6
" 4	78.7	81.1	82.4	5
Names.....Group 1	32.7	41.5	41.4	27
" 2	34.7	39.9	42.7	23
" 3	35.3	39.7	42.1	19
" 4	35.5	41.5	45.9	29

Average % of gain of Group 1 in all tests........................ 13.0
" " " " " 2 " " 16.3
" " " " " " 3 " " " 15.6
" " " " " " 4 " " " 17.0

Sleight, by his method of computation, found only a few instances of significant amounts of transfer. His conclusion is that "There appears to be no general memory improvement as a result of practice, nor any evidence for the hypothesis of a general memory function" (p. 455).

After the conclusion of these experiments, Sleight repeated the same investigation with some modifications, on a group of young women, 18 to 19 years old. The results were substantially the same.

Coover made a study of the effect of training in reproducing imagery of a simple kind upon ability to reproduce imagery aroused by materials of various sorts. The tests made before and after the training were as follows: (1) recognition or choice of one of two letters previously shown, (2) reproduction and recognition of letters presented in groups of 12, (3) discrimination of intensities of sounds, (4) memory of visual symbols. The training consisted in practice in discriminating intensities of sound, and extended through a period of 48 days. These intensities of sound were produced with a sound pendulum (wood) instead of with a fall phonometer (steel) as in end test number (3). The results of the investigation show small or doubtful effects of transfer.

"The training on discrimination of sound did not result in improvement in efficiency with the training material. But, according to introspective evidence, it effected changes in the processes employed. Quantitative analysis showed that the practice-effect of the evident exercise of retention and reproduction of auditory and other imagery 'spread' to the tachistoscopic test of Recognition or Choice of One of Two Letters, and to the test on the Complete Learning of series of visual symbols, both of which involved retention and reproduction of imagery."

Dearborn made some experiments to measure the effect of practice in learning vocabulary and poetry upon ability to memorize various sorts of material as specified in the following table. He did not make the end tests on a control group and hence it is impossible to determine how much of the gain in the end tests was due to the practice series. Judging from other experiments these gains would have to be reduced by one-half or one-third. An interesting comparison may be made between the gain in the end tests and the training series. The average gain in learning French and German vocabulary was 57%, whereas the average gain in the end tests in learning French, German, or English verse was only 19% or one-third as much. Practice in learning *Paradise Lost* made no improvement in learning chemical formulæ.

TABLE 54

Transference of practice in memorizing German and French vocabularies and English poetry and prose. After Dearborn (1910), p. 385

Person	Practice Material	Percentage Gain	End Test Material	Percentage Gain
1	French Vocab.	57%	French Verse	25%
2	German "	60	German "	10
3	French "	53	English "	17
4	" "	55	" "	7
5	" "	62	French "	33
6	German "	57	German "	25
7	Victor Hugo	82	Browning	52
8	Horace's Odes	73	Norse Poem	17
9	Paradise Lost	68	Chemical Formulæ	0
10	Enoch Arden	55	Burke	2

Bennett ('07) had one person memorize 16 lines of *In Memoriam* a day for 28 consecutive days. This person was tested before and after this period of training by learning a list of 15 names of places each day for five days in which he showed a gain of 58 %. Another person memorized two stanzas of *Faerie Queene* a day for 35 consecutive days. Before and after this period he was tested in learning a list of 30 digits each day for five days and in which he showed a gain of 22%. No control tests were made on other persons without training in learning the poetry.

Winch tested a group of 34 girls, averaging 13 years of age, by having them learn a passage of historical prose. On the basis of this test he divided them into two groups of equal ability. Group A memorized 18 to 20 lines of poetry each day on four days scattered through a period of two weeks. Group B meanwhile worked sums. At the end of that time both groups were tested with historical prose. Group A rose from a total score of 1,497 to 2,055, or 37%, while group B rose from 1,497 to 1,890, or 27%.

Winch ('08 and '10) next tested another class of 34 girls in the same general manner, except that the before and after test was made with geographical passages and that the poetry for the training series was somewhat simpler. He also carried out a similar experiment with a third class of girls, using a historical passage for the end tests. The results in each case showed a greater gain in the practiced group.

The results from the author's class experiments (Chapter XI, *Experiments*) indicate that an improvement of 27% in learning Italian vocabulary is accompanied by an improvement of only 8% in learning French vocabulary or less than one-third as much. Improvement in transcribing letters into numbers was accompanied by only 12% as much gain in transcribing numbers into symbols.

Attention. Coover made a series of tests on transference which he lists under the head of attention, but it is doubtful, as he himself states, whether they are measures of attention any more than many other tests that have been reported under other headings. At any rate, attention probably played an important part in most of the tests that Coover employed. The following were the nineteen end tests made before and after training.

I. Reaction		
1. Simple sensory to sound	(50)	1
2. Compound		
a. With discrimination		
(1) Marking out small a's	(100)	2
(2) Marking out o's	(100)	3
b. With discrimination and choice		
(1) Card-sorting	(200)	4
(2) Typewriter-reaction	(200)	5
(3) Controlled reaction	(50)	6
II. Sensible discrimination of sounds	(90)	7
III. Reproduction		
1. Unequivocal (Rote memory)		
a. Successive presentation		
(1) Memory of sound intensities	(50)	8
(2) Memory of consonants	(50)	9
(3) Memory of Arabic numerals	(50)	10
(4) Memory of visual signs	(10)	11
(5) Memory of associated pairs	(50)	12
b. Simultaneous presentation		
(1) Learning 12-letter-rectangles		
(a) Free	(10)	13
(b) With distraction	(10)	14
2. Equivocal—Word-completion	(10)	15
3. Free—2-minute trains of ideas	(3)	16
IV. Extensive threshold of visual attention		
1. Free	(15)	17
2. With distraction	(10)	18
V. Maximum voluntary activity—tapping	(5 30″)	19

(The figures in parenthesis indicate the number of reactions, memory units or experiments, in the test.)

"These tests were taken by 10 reagents, 8 of whom took training between the first and final series which were separated by an interval of

55 days. The first series of tests occupied 12 days during a period of 36 days; the second, or final, 10 days during a period of 21 days. Each pair of tests was separated by an interval of about 66 days.

"The pairs of tests were also taken by two reagents of a group of 21 control reagents. There were thus two sets of control reagents: The two who took all the tests, and the 21 each of whom took only one pair or a few pairs of tests." (Coover, '16.)

Different ones of the subjects took training in different material:

"During the 55-day interval between the tests, two reagents (MN., Le.) took training 18 days on Test 17; 25 12-letter-rectangles were presented daily, making in all 450 experiments each. Two reagents (Rt. and Sl.) took training 18 days on Test 13; 20 12-letter-rectangles were presented daily, aggregating 360 experiments each. One reagent (Ly.) took training in simple reaction to sound for 11 days, 1,100 reactions in all. (Le., who took training on test 17, also took training in this simple reaction to the extent of about 500 reactions.) Two reagents (He., Cr.) took training on memory schemes for about 14 days. And one reagent (al.) took training on Test 17 for 8 days, almost consecutive, to the extent of 200 experiments."

The results obtained from the various end tests are rather intricate and difficult to present in tabular form and somewhat doubtful as to their meaning so far as improvement in attention is concerned. Coover attempted to interpret their meaning from the standpoint of control of attention by comparing the variability in the performance of the persons before and after training on the assumption that reduction in variability indicated better attention. For a detailed consideration, Coover's original report must be consulted. His general conclusion was that "as a measure of attention our tests are inadequate, and the question of transference of improved conditions of attention remains open" (page 183).

g. Analysis and Ingenuity. Ruger ('10), in connection with his study of learning to solve puzzles, made observations on transfer of practice. His results are difficult to summarize in brief form. It will have to suffice, therefore, to say that he enumerates general factors of transfer in solving puzzles as follows: (a) The ideal of efficiency, that is, "the active search for methods of control;" (b) a high level of attention was a precondition of success; (c) attitudes—"The change from the self-conscious to the problem-attitude occurred sometimes automatically, and sometimes deliberately by means of an ideal. The most powerful stimulus to change of attitude and so of its transfer was personal success; it did not matter much whether it was accidental or planned;" (d) methods of attack. As to special factors he mentions: (a) Related ideas—"Geometrical concepts played an almost negligible part in the work of solution;" "The greatest transfer in the way of related ideas was that from similar puzzles;" (b) motor habits—"The mere presence, in the case of change of conditions, of motor habits appropriate to the new conditions did not necessitate positive transfer," "The degree of positive transfer varied directly with the precision of analysis of the similarity of the new case to the old," "In some cases a generalized formula developed in connection with the first case was essential to effective transfer of motor habits to later modifications of the first case," "Transfer was more effective in those cases where the formula or general rule was developed in the first few trials, and where the formation of perceptual-motor habits had been controlled and inter-penetrated by it from the start, than when the generalization had been arrived at after those habits had been set up."

h. Cross Education. Cross education refers to the transfer of practice from one organ of the body to bilaterally symmetrical organs, as for example the spread of training from the right hand to the left hand. A number of investigations have been made on this problem which show that such transfer takes place to a very great extent. Scripture, Smith, and Brown ('94) state that improvement in the strength of grip with one hand produced 80% as much gain in the other. They also report that Volkmann found that improvement in discrimination with the left arm was accompanied by approximately 80% as much gain in the other arm and that other instances showed similar gains.

Davis ('98–'00) measured the effect of practice in tapping with the right great toe upon the rate of tapping with the right hand, the left hand and the left great toe. He found that the left toe improved 151% as much as the right toe, with which the practicing had been done, the right hand 100% as much, and the left hand 83% as much. He also found that practice in gripping a dynamometer with the one hand improved the other about 70% as much. Practice in hitting a target 100 times with the right hand improved the left hand about 75% as much.

Woodworth ('99) reports that practice in hitting dots with the left hand improved the right hand about 50% as much and Swift found that practice in tossing balls with the right hand caused the left hand afterwards to improve in the same exercise more rapidly than it would otherwise have done.

The writer ('10) measured the amount of transfer of improvement in tracing with the right hand a star outline as seen in a mirror to tracing the same outline with the left hand. The test was made by having a tracing made first with the left hand, then a series of 25 to 100 tracings with the right hand, and, at the close, a tracing again with the left hand. The amount of transfer to the left hand is approximately 90% of the total amount of gain made by the right hand. The left hand improves nearly as much as the right hand although all the practicing had been done by the right hand. The results of experiments in cross education are somewhat uncertain in their meaning so far as transfer of training is concerned. Improvement in one organ, which is uniformly accompanied by a very large improvement in bilaterally symmetrical organs, is probably due to the fact that many common processes are involved in doing a task with two bilateral organs. For example, in practicing with the hands, many of the same sensory and neural processes are involved. Thus, the same visual processes and the same visual brain centers would be concerned. Likewise, it is also probable that neural innervations going to the right hand in practice also go to the left hand. These would tend to improve the control of the left hand without actual practice with the left hand. The data on cross education probably have only a distant and doubtful bearing upon the problem of transfer.

Criticism of the Technique of Experiments on Transfer of Training. There are three important elements in the technique of experimentation in this field which have not been recognized by investigators from the beginning and are not recognized by all investigators even to-day. (1) The first is the length of the end tests. These have been too long in some investigations to give as full opportunity as possible to transference. This was one of the difficulties in James' original experiment and was recognized by James himself. (2) In the second place, the end tests have not always been repeated on a control group of subjects. This is true of nearly all of the early studies. (3) In the third place, many investigations do not mention the amount of improvement made in the training series itself with which the gain in the end tests may be compared. Failure to observe these precautions makes impossible an accurate, quantitative interpretation of many of the early researches and even of some of the recent ones.

Summary. It may perhaps be unwise in view of the intricacy of the researches and their partial incongruity to attempt to summarize general conclusions. However, a brief resumé will help to clarify the reader's thinking about these problems. (1) Practically every investigation shows that improvement in one mental or neural function is accompanied by a greater or less amount of modification in other functions. (2) This modification is in most instances a positive transfer, that is, an improvement. Negative transfer, that is, loss of efficiency in other functions, or interference, has been reported principally among sensori-motor habits. (3) The amount of improvement in the capacity trained is probably never accompanied by an equal amount of improvement in other capacities, with the possible exception of a few isolated instances whose actuality may be questioned. Thus, for example, Thorndike and Woodworth found that the gain in various types of perception or discrimination closely related to the type in which the training took place was from 0% to about 40% as great as that made in the particular kind of perception trained. In memory, Fracker's results showed that the improvement in different sorts of memorizing, so similar to the training series that they were all but identical with it, was about 75% as much as that made in the training series; while the improvement made in the forms of memory rather different from the training series was only about 15% as much as that in the training series. Up to about 1890 when James reported the first investigation on the problem of transference, it was tacitly assumed by many writers that a very large share, if not all, of the training derived from one sort of exercise was carried over to other sorts of exercise. After the first investigations became generally known, many writers went to the other extreme and assumed that all training is entirely specialized and that nothing carries over from one kind of practice to any other kind of practice. As a general estimate, on the basis of experimental work done thus far, the amount of transference between the extremes of 100% and 0% of transfer lies nearer to the zero end and is probably in the neighborhood of 20% to 30% of transfer to closely allied functions and from that point on down to 0% of transference to more unlike functions. (4) In the fourth place, the improvement spread to other functions diminishes very rapidly in amount as these other functions become more and more unlike the function specifically trained. This diminution occurs at a surprisingly rapid rate.

How Does the Transfer Take Place? If improvement in one mental function is accompanied by, or produces improvement in, other functions, how may the change in these other mental functions be explained? How does change in one function carry over to others? Two general theories have been proposed: (1) The theory of identical elements or special connections, and (2) the theory of generalization or common capacities.

The theory of identical elements has been advocated by Thorndike and may best be stated in his own words:

"The answer which I shall try to defend is that a change in one function alters any other only in so far as the two functions have as factors identical elements. The change in the second function is in amount that due to the change in the elements common to it and the first. The change is simply the necessary result upon the second function of the alteration of those of its factors which were elements of the first function, and so were altered by its training. To take a concrete example, improvement in addition will alter one's ability in multiplication because addition is absolutely identical with a part of multiplication and because certain other processes,—e. g., eye movements and the inhibition of all save arithmetical impulses,—are in part common to the two functions.

"Chief amongst such identical elements of practical importance in education are associations including ideas about aims and ideas of method and general principles, and associations involving elementary facts of experience such as length, color, number, which are repeated again and again in differing combinations.

"By identical elements are meant mental processes which have the same cell action in the brain as their physical correlate. It is of course often not possible to tell just what features of two mental abilities are thus identical. But, as we shall see, there is rarely much trouble in reaching an approximate decision in those cases where training is of practical importance." (Thorndike,'14, II, pp. 358–359.)

The theory of generalization has been advocated by Judd in the following manner:

"The important psychological fact . . . is that the extent to which a student generalizes his training is itself a measure of the degree to which he has secured from any course the highest form of training. One of the major characteristics of human intelligence is to be defined by calling attention, as was pointed out in the chapter on science, to the fact that a human being is able to generalize his experience. James has discussed this matter by using the example of the animal trained to open a particular latch. The animal becomes acquainted with the necessary movements to open one door, but he never has the ability to generalize this experience. He cannot see that the same method of opening doors is applicable to many other latches. The result is that the animal goes through life with one particular narrow mode of behavior, and exhibits his lack of intelligence by his inability to carry this single type of skill over to the other cases which are very familiar to the trained human intelligence.

"James goes on to say that the same distinction appears when we contrast a trained scientific mind with the ordinary mind. The ordinary thinker does not see how to deal with a situation in terms of scientific principles. James cites the example of his own experience with a smoking student-lamp. He discovered by accident that the lamp would not smoke if he put something under the chimney so as to increase the air current, but he did not realize that what he had done was only one particular example of the general principle that combusion is favored by a large supply of oxygen. The general principle and its useful application belong to a sphere of thinking and experience which the untrained layman has not yet mastered." (Judd, '15, pp. 413-414.)

The theory of identical elements is based on the doctrine that learning or changes in mental capacities consist of the establishment of specific connections or associations between various specific elements. One form of exercise has influence upon another capacity whenever connections established in the former may also be used in the latter. In a certain sense the theory of identical elements describes or explains transfer of training in a tangible, concrete manner. In a certain other sense it does not explain transfer of training at all or else it implies that there is no general training in the sense in which formal disciplinarians use the term. If special training is general or helps in performing various mental activities only to the extent to which the special training has elements in it which occur also in these other activities then there is no spread of training to such activities in which no elements are found which also appear in the capacity specifically trained. The formal disciplinarian assumes that training of one sort affects capacities of other sorts irrespective of identical elements or similarity to the activities developed. In the last analysis the controversy comes down to a question of fact, namely, to how dissimilar activities does any given form of training spread? The theory of identical elements, when the term identical elements is used in a liberal manner, has the advantage of describing the situation in concrete, definite concepts and lends itself fairly well to the interpretation of experimental results. The discussion of the formal disciplinarian is usually not in as tangible terms but is likely not to be very different from the statements of the experimentalist when the former reduces his argument to specific terms.

The theory of generalization attempts to explain spread of improvement in terms of the recognition of application of an experience obtained in one connection to other connections and is probably more satisfactory to the formal disciplinarian. In the author's opinion there is no necessary opposition between the theory of identical elements and the theory of generalization. The essential difference is in the emphasis upon the conscious recognition of identical elements in as many situations as possible. Judd has emphasized this in connection with teaching:

"The first and most striking fact which is to be drawn from school experience is that one and the same subject-matter may be employed with one and the same student with wholly different effects, according to the mode of presentation. If the lesson is presented in one fashion it will produce a very large transfer; whereas if it is presented in an entirely different fashion it will be utterly barren of results for other phases of mental life. It is quite possible to take one of the objects of nature study, for example, and to teach it in such a way that it becomes an isolàted and utterly formal possession of the student. This has been illustrated time and time again by the instruction which has been given in birds and plants. A teacher can teach birds and plants in such a way as to arouse a minimum of ideas in the student's mind. The training may be as formal in these content subjects as it ever was in language instruction. On the other hand, the same subject-matter may be taken by a different teacher, and under other methods can be made vital for the student's whole thinking. Thus the teacher who is dealing with birds as a subject of nature study and secures an interest on the part of his students for the world in which these birds live, through an examination of the structures and habits of the birds, will have in this subject-matter one of the most broadly interesting topics that can be taught. In exactly the same way a teacher who knows how to make use of the materials given in a Latin course may render this subject very broadly productive, as contrasted with the teacher who merely gives the formal aspects of the subject. Formalism and lack of transfer turn out to be not characteristics of subjects of instruction, but rather products of the mode of instruction in these subjects." (Judd, '15, pp. 412–413.)

It seems then that the two theories are not necessarily antagonistic but when sanely interpreted are useful supplements to each other. The theory of identical elements has helped to make the discussion of formal discipline or transfer of training concrete, and the theory of generalization will help to emphasize the conscious recognition of the identical elements in as many situations as possible. Some writers have assumed that transfer is limited to a conscious recognition of elements. This, however, is disproved by some experiments with human beings and particularly by the experiments with animals such as those reported by Webb.

CHAPTER XIV

TRANSFERENCE OF TRAINING IN ABILITIES IN SCHOOL SUBJECTS

To what extent does the training of the capacities exercised by school subjects carry over to capacities concerned in other school subjects, and especially to the capacities involved in the usual activities of life? This question brings the problem of transference of training directly face to face with the issues of education and is the form in which it is usually concerned in discussions of mental discipline. It is of more special concern to the liberal phases of education in the high school and the college for the reason that the subjects taught in the elementary school, in the vocational courses in the high school, and in the professional courses in the university are directly pertinent to the common needs of life or to the various occupations and professions. Most of the discussion has, therefore, centered about the training value to be derived from the traditional academic work of the high school and the college.

General Opinions. The beliefs concerning transfer of training of the capacities employed in school subjects have been largely matters of opinion and not matters of fact. These opinions, held by persons prominent in educational affairs, have been for the most part rather uniformly confident in the faith that the exercise of the mind upon the materials of the school subjects produces a very profound improvement in mental powers as a whole. Typical of such opinions are the following:

"But my opinions of the supreme educational value of the great disciplinary studies have not changed, and will never change.

"As a result of my long experience in watching their effects on our students I am absolutely and irrevocably sure that certain subjects train in thinking straight and reasoning clearly.

"I am absolutely sure that Latin and Greek, higher mathematics, philosophy, the critical study of the literatures of different nations (and the better the literature, the better the training it gives, Greek, Latin, English, and French literatures leading all others in this respect, and in the order named), economics and politics, especially on their theoretical side, and English composition are thinking subjects of very high educational value." (Thomas, M. C., "Old Fashioned Disciplines," *Journal of the Association of Collegiate Alumnæ*, May, 1917, p. 588.)

In connection with the Classical Conference at Ann Arbor, Michigan, 1909, Dr. Harvey Wiley sent out questions to 100 scientific men asking what they considered to be the value of Latin and Greek as preparation for scientific pursuits. He received 35 replies distributed as follows:

Favorable to the study of Latin and Greek	14
Unfavorable to the study of Latin and Greek	17
Favorable to Latin, but not to Greek	4

Among the opinions expressed in this connection, Professor R. P. Biglow made this statement:

"To summarize my opinions in the matter of a scientific education, it seems to me that the essentials are of two classes: First, a thorough training in the use of the tools required by a scientific man, namely, the modern languages and mathematics; second, a training in the scientific method, especially as applied to the branch of science in which he desires to specialize. If to the curriculum, the studies of classics can be added without interfering with these essentials, then it seems to me that in some cases it would be desirable as a means of culture."

Professor Neff of the University of Chicago regretted the time he spent on Latin and Greek:

"I think everyone realizes as he grows older that he has his limitations. I, for one, regret very keenly that I took a great deal of Latin and Greek and did not spend far more time on advanced mathematics and physics. I am, however, not now wasting any time in vain or useless regrets on this account, but simply doing the best I can with the knowledge that I have acquired."

The opinions of prominent business men were reported at the same conference:

Mr. William Sloane, a New York business man:

"I believe that the slow processes of translation of the classics make good training for the boy who has chosen a business career."

The Hon. J. W. Foster, of Washington:

"The mere routine labors of the translation of Greek and Latin authors into one's vernacular, the effort to ascertain their exact meaning and the choice of the words which correctly express that meaning, constitute a mental training which will be invaluable to the future lawyer or public man."

Probably the most notable assemblage of opinions ever brought together was that presented at the conference on classical studies held at Princeton University in June, 1917.[1] These opinions were quoted from some 300 persons prominent in public life, business, universities and colleges, schools, the ministry, law, medicine, engineering, science, journalism, modern literature, history and related branches, fine arts, and oriental studies. These statements were substantially unanimous in bearing testimony to the value of the classical languages.

But perhaps as many opinions on the other side from men of equal intellect could be gathered. The point is that mere opinions cannot yield a final appraisal of the training value of school studies. Both favorable and unfavorable opinions are bound to be very nearly worthless because at best they are apt to be prejudiced by personal likes or dislikes and by exceptional instances of benefit or lack of benefit from the pursuit of this or that particular subject, and most of all because no general observer has at hand sufficient, reliable or complete evidence concerning the problem. Experimental and statistical data are hard enough to interpret because of the complication of factors in the production of any type of training to say nothing of the settlement of the controversy by general impressions.

We cannot determine by ballot the shape of the earth, or the value of a patent medicine, no matter how many testimonials may be presented on the one side or the other. Men prominent in life have testified to the benefits of patented remedies which science has shown to be not only valueless but harmful.

Specific Estimates of the Value of School Studies. The writer attempted to obtain specific estimates of the value of school subjects according to the best judgment that could be exercised by persons who are concerned with the work of public education. These were obtained not with a view to contributing anything toward the solution of the problem but for the purpose of examining more precisely the drift of the consensus of opinions held by persons immediately in charge of school work.

Any branch of learning may have three possible values—a

[1] Reported in *Value of the Classics*, Princeton University Press.

disciplinary value, a utility value, and a cultural value. Thus the study of English has a certain amount of disciplinary value in training the mental capacities involved in the learning and understanding of the material studied; it has an obvious utility value in acquiring the ability to use English correctly and effectively; and it has a certain cultural value in acquainting the student with the thought and life of mankind.

Estimates of the three values of each of the subjects listed in the table were made as carefully as possible by fifty-eight superintendents, principals and teachers. In making these estimates it was assumed that the pure, disciplinary value of the first year of high school English as taught in the average way be equal to 10 and that all other values be estimated in terms of this assumption. If the disciplinary value of algebra was considered to be twice as great as that of English, hour for hour devoted to each, then it should be estimated as 20. Or if the utility or cultural value of English was considered greater or less than its disciplinary value, the rating should be indicated accordingly. It was further assumed that these were to be the values for the average boy or girl in the high school.

At first glance it would seem that such judgments would be rather uncertain and variable, and, as a matter of fact, they were quite variable. Nevertheless, viewed from the statistical standpoint, the judgments present a normal distribution spreading over a wide range but clinging in large numbers about a central point. For example, the judgments of the disciplinary value of American history ranged as extremes from 3 to 30, with the largest number of estimates on 10 and a gradual decrease in the number of estimates on values farther and farther removed from 10. The median estimate was 10 and the probable error was 3.5, that is, one-half of the estimates were between 8 and 15. So that the judgments about the various values, even though quite variable and usually accompanied by a feeling of uncertainty so far as the individual judge was concerned, were as reliable and as normal as judgments about most matters are.

The following table gives the median judgment for each of the values listed:

TABLE 55

Estimated values of school subjects. After Starch ('17)

	DISCIPLINARY VALUE	UTILITY VALUE	CULTURAL VALUE	TOTALS
Geometry	20	10	8	38
Algebra	19	9	8	36
Latin	17	10	14	41
Physics	17	18	12	47
Gymnastics	17	12	7	36
German	15	15	12	42
French	13	11	13	37
Chemistry	13	19	11	43
Manual Training	12	23	10	45
Football	12	6	5	23
Shorthand	11	21	5	37
English (assumed)	10	30	22	62
U. S. History	10	18	17	45
Physical Geography	10	15	10	35
Music	10	11	25	46
Cooking	10	30	9	49
Bookkeeping	10	25	7	42
Civics	10	21	12	43
Botany	10	12	11	33
Zoölogy	10	11	11	32
Drawing	10	11	15	36
Sewing	9	30	10	49
Typewriting	9	22	5	36
Work of teacher or business man	22	33	18	73
Earning one's way through school	24	36	18	78

An examination of the table reveals some interesting comparisons. The highest disciplinary value is assigned to geometry with a rating of 20, or twice as high as that of English. Algebra is next with a value of 19. It is rather surprising to find gymnastics and football rated as high as they are. It is also interesting to note that the disciplinary value of a pupil's earning his way through school is rated higher than that of any of his studies. The lowest disciplinary value is assigned to sewing and typewriting.

In the case of utility values, the highest rating is given to English, cooking and sewing (30), a value approximately three times as great as their disciplinary value. The lowest utility value is assigned to football (6) and the next to algebra (9). Again the utility value of a pupil's earning his way through school (36) is placed above that of any of his studies.

In the case of cultural values, the highest rating is given to music (25) and the next to English (22). The lowest value is assigned to shorthand, typewriting, and football (5). The cultural value of a pupil's earning his way through school (18) is placed below only music and English.

A similar study with reference only to the disciplinary aspect of college studies was previously reported by Thorndike ('15). Estimates obtained from 100 teachers are summarized thus:

"Philosophy (for freshmen) 8; English composition, 10; German, Chemistry and Logic, 11; Physics, 13; Latin, Greek, Mathematics, 16. Waiting on table is rated at 3; athletics is rated at about 7; work for the college paper at 8 or 8½; tutoring at 13 or 13½; and regular productive work in the world as teacher, business man or skilled laborer at 14½." (Page 281.)

Such tabulations of opinions are valuable only in showing in more accurate terms what teachers and educators think about the question and not in really answering it. However, so long as schools are operated by opinions, a combination of opinions may be better than individual ones as guides of educational policies. To what extent individual opinions are consciously or unconsciously prejudiced by personal interests is shown by the fact that the teachers overestimated by nearly one-half the value of their own specialties as compared with the average values assigned by teachers as a whole.

Experimental and Statistical Inquiries. *a. Arithmetic.* The writer made an investigation to measure the effect of improvement in mental multiplication of three-place numbers by a one-place number, doing 50 problems a day for 14 days, upon other types of arithmetical processes. This experiment was carried out with eight subjects who constituted the training group, and seven subjects who constituted the control group. The results are given in the following table, which shows the percentages of gain of the second end tests over the first:

TABLE 56. After Starch ('11)

	Trained Persons	Untrained Persons	Differences
Adding fractions	40	12	28
Adding three-place numbers	49	10	39
Memory span for numbers	−3	−2	−1
Subtracting numbers	58	35	23
Multiplying four-place numbers	53	29	24
Memory span for words	3	−5	8
Multiplying two-place numbers	47	10	37
Dividing three-place numbers	45	25	20
	—	—	—
Averages, exclusive of memory span	49	20	29

The residual gain on the average was 29%. The average gain made by the trained group in the practice series, comparing the first day with the 14th day, was 112%. Hence the gain transferred to the allied arithmetical operations was only 26% of the gain in the practice series itself. From one point of view, this seems to be a very considerable amount of transfer, but when we note that some of the end tests were as similar to the training series as they could be without being identical with it the transfer is small. We might expect almost a complete carrying over to the closely similar operations but the largest amount of residual gain took place in the multiplying of two-place numbers by a one-place number and in the adding of three-place numbers; but even there it was only slightly larger than the transfer to the other operations.

Winch conducted a series of experiments to determine the amount of transfer from improvement in numerical computation to arithmetical reasoning. In each experiment the class was divided into two groups of approximately equal ability as shown by a previous test in arithmetical reasoning. Then one-half of the class was trained in "rule" sums after which a final test in arithmetical reasoning was given alike to both groups.

The first class, composed of 13-year-old girls from a poor neighborhood, showed improvement in numerical accuracy but no transfer to arithmetical reasoning. The second class, composed of 10-year-old girls from a poor neighborhood, showed considerable improvement in accuracy and a doubtful transfer to reasoning. The third class, composed of 10-year-old girls from a good neighborhood but poor in arithmetic, showed transfer in three sections but the opposite in a fourth section. In the fourth class, composed of 10-year-old boys, both practiced and unpracticed groups showed about equal gains in arithmetical reasoning. Winch concludes from these experiments that improvement may take place in numerical computation without any certainty of improvement in arithmetical reasoning.

In a later experiment conducted on the same general plan, he used 72 ten-year-old boys. Here again he found no evidence of transfer from improvement in numerical computation to arithmetical reasoning.

Carrie R. Squire made an experiment regarding the transfer of neatness: "At the Montana State Normal College careful experiments were undertaken to determine whether the habit of producing neat papers in arithmetic will function in reference to neat written work in other studies; the tests were confined to the intermediate grades. The results are almost startling in their failure to show the slightest improvement in language and spelling papers although the improvement in arithmetic papers was noticeable from the first." (Bagley, *The Educative Process*, p. 208.)

This experiment was repeated under the direction of Ruediger with the difference that along with the specific training in neatness in one particular study a general practice of neatness in daily life was held up before the class as an ideal to be striven for. Care was taken not to discuss neatness in the other classes. Sample papers were taken in the one subject concerned and in two other subjects before and after the training. The seventh grades of three schools, located at widely different places and comprising 39 pupils in all, were used in the experiment. The two schools which showed an appreciable improvement in neatness in the study where special training was given also showed considerable, though less, improvement in the other studies. Thus the two schools showed an average improvement of 4.75 points in the study where training was given and 3.1 points, or 65% as much, in studies where nothing was said about neatness.

b. Grammar. Grammar has been regarded as a highly efficacious instrument for training the functions of the mind. Thus Commenius stated: "I presume that no one can raise any objection to my placing (Latin) grammar first, since it is the key of all knowledge." Locke said on the other hand: "I would fain have anyone name to me that Tongue, that anyone can learn, or speak as he should do, by the rules of Grammar. Languages were made not by Rules, or Art, but by Accident, and the Common Use of the People."

Claims for Grammar have been that it

1. Disciplines the mind.
2. Prepares for the study of other languages.
3. Gives command of an indispensable terminology.
4. Enables one to use better English.
5. Aids in the interpretation of literatures.

The Committee of Ten (1893) said: "The study of formal grammar is valuable as training in thought, but has only an indirect bearing on the art of writing and speaking."

What are the actual facts so far as any are available at the present time? Briggs ('13) attempted to determine the extent to which the various claims made for grammar are substantiated. He outlined the following claims and devised an elaborate set of tests to measure the effects of training in grammar.

"It is held that grammar trains children:

A. With rules and definitions:
  1. To see likenesses and differences.
  2. To critically test a definition.
  3. To thoroughly apply a definition.
  4. To make a rule or definition.

B. With reasoning:
  5. To test reasons.
  6. a. To take from a mass of data all that are necessary and to use them in reaching a judgment.
     b. To demand all necessary data before drawing a conclusion.
  7. To reason in other fields, e. g., arithmetic.
  8. To reason syllogistically.
  9. To detect "catches."

As illustrations of the nature of the tests we may cite the following instances. For measuring the observation of likenesses and differences, Briggs used such a test as this:

"One-half of the following 16 words are alike in one respect and in that respect unlike all the others in the list. Find these eight words and mark them with a check (√)."

biscuit	pirate	mountains	men
oxen	geese	fathers-in-law	factory
scholars	knives	vessel	table
pole	frame	children	mice

(8 are plurals)

To determine ability to judge definitions and to amend them, he used as tests such statements as the following for shoe:

1. A portion of clothing.
2. Something black made of leather.
3. Something to wear on the feet.
4. A necessary article costing from $1.00 to $5.00 or $6.00.

These tests were given to 25 or 30 pupils in each grade from two to seven in the Horace Mann school. Each class was divided into two divisions. Then for three months, three times a week, the children of Division I were taught formal grammar. During the same three months, the children of Division II had work in composition and language. They were then given the second set of tests similar to the preliminary tests, after which the conditions were reversed. Division II then had formal grammar and Division I had language and composition work. At the conclusion of this period, the first set of tests was again given to all of the children. The upshot of the whole investigation is summarized by Briggs in the following manner:

"As a result of this experiment it may safely be asserted that these particular children after the amount of formal grammar that they had, do not, as measured by the means employed, show in any of the abilities tested, improvement that may be attributed to their training in formal grammar."

Hoyt ('06) made a study to determine the relation between the knowledge of grammar, ability to interpret English, and ability to write English. He employed three tests: One for grammar consisting of ten questions on four stanzas of Gray's *Elegy;* the second for testing ability to interpret English, consisting of a statement of thought in four other stanzas of Gray's *Elegy;* and the third for ascertaining ability in composition, consisting of writing a composition in forty minutes. These tests were made with 200 pupils in a high school in Indianapolis. All papers were marked by two examiners according to the percentage method. Correlations were then computed among the different tests, which were as follows:

Grammar and composition	.23
Grammar and interpretation	.28
Interpretation and composition	.32

These coefficients are very low and indicate that a greater or less amount of knowledge of grammar is accompanied to only a slight extent by greater or less ability to write a composition. The same statement holds for the relations between the other comparisons of abilities. The fact that the pupil who knows more or less grammar writes respectively a slightly better or worse composition is quite likely due to the fact that he is a better or poorer pupil rather than to any aid which knowledge of grammar may render him in writing a composition.

Hoyt concludes that " . . . the teaching of grammar is of little avail in strengthening one's ability to use language."

The writer ('15) made a series of tests in formal grammar and in correctness of English usage. The test in formal grammar consisted of three parts: First, a passage in which the parts of speech of as many successive words as possible were to be indicated in three minutes; second, a passage in which the cases of nouns and pronouns were to be indicated in three minutes; and third, a passage in which the tenses and modes were to be indicated in three minutes. The test for usage consisted of 100 sentences each of which was stated in two ways. Both might be correct, both might be incorrect, or one might be correct and the other incorrect. Pupils were allowed fifteen minutes in which to indicate the correct expressions. The results are summarized in the following tables.

The tests were made upon 54 university Juniors and Seniors and 146 high-school pupils. They gave the results shown in Table 57, in which the scores for knowledge of grammar are the numbers of the parts of speech, tenses, cases, and modes indicated correctly in the specified period of time, and the scores for correctness of usage are the numbers of sentences designated correctly in the specified period of time.

TABLE 57. After Starch ('15)

Years of Foreign Languages	Number of Students	Average Scores for Knowledge of Grammar	Average Scores for Correctness of Usage
	University Juniors and Seniors		
0	2	48.0	81.5
2—5	12	47.8	71.1
6—9	25	58.6	75.5
10—15	15	63.4	75.7
	High School Pupils		
0	12	14.7	32.2
8 weeks	50	20.8	43.0
1 year	18	25.5	43.4
2 years	39	24.8	45.9
3 years	27	28.6	47.7
	University Juniors and Seniors		
Years of Latin			
0	15	45.8	70.9
1—3	11	56.1	75.7
4	14	57.5	74.3
5 or more	9	51.8	76.1

Another test for correctness of usage, consisting of sentences like the set of one hundred but arranged in the order of increasingly difficult steps, was made on another group of 146 university students and 92 high-school pupils. This test yielded the results given in the following table. The scores are the numbers of the highest steps passed. The higher the score is, the greater is the ability of using English correctly.

TABLE 58

Years of Latin	Number of Pupils	Average Scores
	University Students	
0	47	10.1
1–6	99	10.2
	High School Pupils	
0	78	9.0
1–4	14	9.3

These tables agree in showing one very significant result, namely, that the study of foreign languages materially increases a pupil's knowledge of English grammar but only slightly increases his ability in the correct usage of the English language. Notice, for example, the upper part of Table 57. The students who had 10 to 15 years of foreign languages made a score in grammatical knowledge of 63, as compared with a score of 47.8 made by the students who had 2 to 5 years of foreign languages, a difference of 32.6% in favor of the former group. For correctness of usage, the corresponding difference is only 6.4%. The two students with no foreign languages made high scores because they were exceptionally good students, but they are too few in number to be considered. The high-school pupils show a gain in grammatical knowledge of 37.5% from the 8-week group to the 3-year group and a gain in usage of only 10.9%. The twelve pupils with no foreign language made low scores because they were exceptionally poor pupils. This is indicated by their low scholarship records, by the fact that many were over-age, by the fact that they avoided the foreign languages, and also by the large difference between their scores and those of the 50 pupils who were just beginning foreign languages. Eight weeks of foreign languages could hardly have produced such a big gain. Their higher scores must be due largely to a difference in original nature.

c. Foreign Languages. Extensive and confident claims have been made for the value of general mental training to be derived from the study of languages. Thus Lodge states the value of the study of Latin as follows:

"Far above every other subject it trains (1) the process of observation, (2) the function of correct record, (3) the reasoning power and general intelligence in correct inference from recorded observation. To this should be added its great value in developing the power of voluntary attention.

"The value of Latin as a practical subject has to do particularly with the effect of the language in the cultivation of English style. In the English vocabulary a very large proportion of words in everyday use are of Latin origin, and it has been estimated that two-thirds of the Latin vocabulary of the classical period has in some form or other come over into English speech. For the correct use of synonyms in English and the habit of expressing one's thoughts clearly, concisely, and cogently, a discriminating knowledge of Latin is indispensable, and while not every pupil in the school may be expected to develop a good style, nevertheless he should be given the necessary foundation for it.

"When we turn to literature, we find that Latin is influential everywhere—particularly in our classical authors—by allusions, by quotations, by actual domestication. Many of our great English writers are permeated with Latin. We cannot expect that all will desire to feed their minds on the works of our greatest authors, however much we might prefer it; but certainly we should not deprive them of one of the most important elements in their enjoyment should they be so minded." (Lodge, p. 388, in *Principles of Secondary Education*, Edited by Paul Monroe.)

Swift measured the progress in learning a new language made by pupils with different amounts of previous language study to determine, if possible, the advantage in beginning a new language to

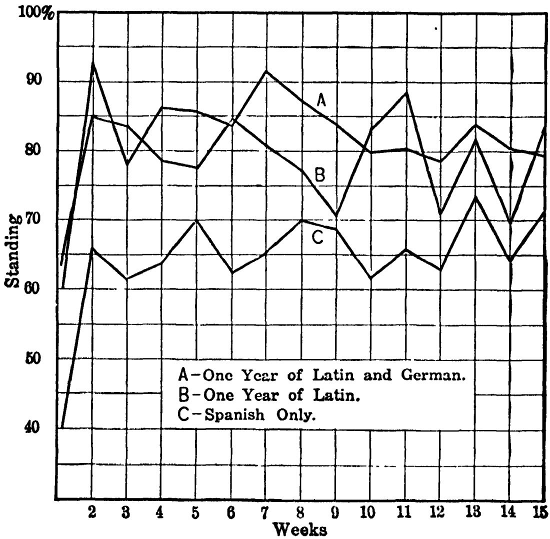

FIG. 55.—Progress of pupils in learning Spanish. After Swift ('06).

be derived from the previous study of another language. The experiment was carried out with two classes, composed of 24 boys and 24 girls in a St. Louis high school, who were beginning the study of Spanish. Weekly tests were made to measure their progress and individual abilities in learning Spanish. The classes were taught in the usual manner and the pupils knew nothing of the purpose of the tests. A record of progress was kept for the first 15 weeks. The results are shown in the following graphs (Figure 55), which indicate the relative progress of the three groups, namely, those who had previously studied one year of Latin and one year of German, those who had only one year of Latin, and those who had never studied a foreign language before. The "Latin" and the "Latin and German" groups stood considerably higher at the beginning than the "Spanish only" group, but the "Spanish only" group gained gradually so that at the end of the 15 weeks it had made up about two-thirds of the difference. Swift concludes:

"The number included in these tests was too small to serve as a basis for anything more than tentative conclusions, but the results certainly open the question whether the advantage to beginners of a new language, so generally thought to accrue from the study of Latin, may not be due, chiefly if not solely, to grammatical information that would be carried over from one language to another, and which would naturally help enormously at the start. In acquiring facility in the use of the Spanish gender, to cite one example, Latin would aid materially, since the majority of Latin feminines are feminine in Spanish, and a large part of Latin masculines and neuters become masculine in Spanish. The declension of Spanish adjectives for gender and number, and their agreement, in these respects, with their nouns, would give Latin students a further advantage. The teacher of the Spanish classes noted that more frequent and detailed explanations of case were needed by those who had not studied Latin. The order of words, also, was more readily mastered by those familiar with the Latin arrangement. Finally, in learning the conjugations and in understanding the significance of tenses, the assistance of the information acquired under these topics in Latin was found to be especially great. The indications, however, are that the higher records made by the Latin and German pupils were the result of the substance of language information obtained from these studies rather than of any so-called 'language' or 'mental discipline.'" (Swift, '06, pp. 250 ff.)

The writer ('15) made a comparison of the scholastic records of university students who had entered the university with two to four years of Latin with the records of those who had entered with two to four years of German. The average grade for the four years of college work of each of the graduates of the College of Letters and Science of the year 1910 was computed. The median mark of the 104 students who had entered the university with Latin was 85.7 and the median mark of the 45 students who had entered with German was 84.0. Hence the difference between the two groups was only 1.7 points.

The explanation for this small advantage of Latin over German may be sought in three directions: First, the disciplinary difference between Latin and German is either zero or very small. Second, whatever difference they may have produced originally may have tended to disappear in the four years of college work, owing to the freedom of electives, pursuit of different courses, disciplinary effect of other studies, etc. Third, the small difference in scholastic records may be due to an original difference in the students themselves, owing to the possibility that one language may attract a better class of pupils than another.

To determine what part, if any, the first two factors played, the average grade of each of the 738 Freshmen of the year 1909–1910 was computed. The median grade of the 416 Freshmen who had entered with Latin was 82.4 and that of the 322 Freshmen who had entered with German was 81.0. Hence the difference between the two groups was only 1.4 points, or approximately the same as that for the graduates.

The next problem was to compare the grades of these two groups in specific subjects as follows:

TABLE 59. After Starch ('15)

Median grade in modern languages of 362 Freshmen who had entered with Latin	84.5
Median grade in modern languages of 293 Freshmen who had entered with German	82.3
Difference in favor of the Latin group	2.2
Median grade in Freshman English of 54 students who had entered with Latin only	83.9
Median grade in Freshman English of 97 students who had entered with German only	82.7
Difference in favor of the Latin group	1.2
Median grade in first-year French of 27 Freshmen who had entered with Latin only	81.5
Median grade in first-year French of 34 Freshmen who had entered with German only	82.0
Difference in favor of the German group	.5

The differences again are very small. The claim of language teachers, so commonly made, that beginners in French who have had Latin are much superior to those who have not had Latin, or that students in English with previous training in Latin are superior to those without such training is ill founded.

Another tabulation was made to show the scholarship records of

Freshmen in relation to the amount of foreign language studied, irrespective of what the languages were.

TABLE 60. After Starch ('15)

Years of Foreign Languages	Number of Students	Median Grade in all Freshman Studies
0	25	81.8
1–2	224	81.9
3–4	195	83.05
5 6	155	84.0

The next problem was to measure the extent to which a pupil's English vocabulary is increased through the study of Latin. The method employed measured the percentage of words of the entire English vocabulary, as well as the approximate absolute number of words, whose meaning a person knows. The test was made with 189 university students and with 46 Juniors in the Madison High School.

TABLE 61. After Starch ('15)

	%
Size of English vocabulary of 139 university students who had studied Latin	60.9
Size of English vocabulary of 50 university students who had not studied Latin	58.2
Size of English vocabulary of 14 high-school Juniors who had studied Latin	54.7
Size of English vocabulary of 32 high-school Juniors who had not studied Latin	50.2

The differences between the Latin and the no-Latin groups are surprisingly small. Nevertheless, the study of Latin does produce an appreciably larger English vocabulary. This advantage becomes less in university students with whom it is partly counterbalanced by the increase in vocabulary due to wider experience.

Partridge compared the standing in the regents' third year English examination of 783 pupils by dividing them into groups according to the number of years of Latin they had studied. His tabulation is as follows:

TABLE 62. After Partridge ('15)

The entire 783 papers divided on a basis of the number of years Latin was studied

Number years studied	0	1	2	3
Number papers written	181	123	220	259
Average standing (percentage)	65	65	69	76

The pupils with no Latin may have had one or more years of other languages and consequently Partridge presents the following table for students who had studied Latin only:

TABLE 63. After Partridge ('15)

Includes papers of pupils having Latin only and no other language to their credit. Total number of papers, 167

No. of years studied	0	1	2	3
No. of papers written	28	25	42	72
Average standing (per cent)	63	61	69	78

TABLE 64. After Partridge ('15)

Includes papers of pupils having German or French only and no other language to their credit. Total number of papers, 176

No. of years studied	0 [1]	1	2	3
No. of papers written	28	41	57	50
Average standing (percentage)	63	61	65	68

Partridge believes that the "superiority of the classical over the non-classical pupils is due not solely to initial natural ability, but to the training received in Latin." He has, however, failed to show the differences in initial ability and consequently any inference of this sort is doubtful.

Harris made a study of the effect of knowledge of Latin upon ability to spell English words by submitting a list of 50 words of Latin origin to 324 freshmen in the University of Illinois. He gives the following table:

TABLE 65. After Harris ('15)

Years of Latin	0	1	2	3	4
No. of students	90	41	95	54	44
Average	82.1	82 4	80.2	81.5	90.1

He further submitted to the same group of students 10 words of Latin origin which were to be defined. This test gave the following result:

TABLE 66. After Harris ('15)

Years of Latin	0	1	2	3	4
No. of students	90	41	95	54	44
Average	30.5	44.2	45.9	53.0	85.3

[1] It is obvious that the 0 columns in these two tables will contain record of exactly the same pupils.

He also compared the grades in Rhetoric of students who had had various amounts of Latin as follows:

TABLE 67. After Harris ('15)

Years of Latin	0	1	2	3	4
No. of students	53	41	66	28	26
Grade	77.2	79.2	79.5	80 6	81.8

Harris concludes: "From these various results the conclusions in so far as these students are concerned, are obvious. In all fields the four-year Latin students showed a marked lead, and in all but the spelling—which I have considered above—there is a steady retrogression although for the practical purposes the one-and-two-year Latin students might be classed together."

The interpretation of these figures is by no means so obvious. Harris has made no allowance for the native superiority of the students with more years of Latin study. In fact, the probability is that, if we may infer from other studies in which such a deduction has been made, a large part of the superiority is due to original nature. Harris's results as they stand prove little or nothing concerning the effect of training in Latin.

F. M. Foster performed a similar experiment at the University of Iowa; 503 freshmen, about equally divided between the sexes, were given a spelling test of forty words of Latin derivation. The results are given in the following table:

TABLE 68. After Foster ('17)

Number of years of Latin	0	1	2	3	4
Average % of errors (girls)	23	28	25	24	17
Average % of errors (boys)	39	37	29	28	27

As in the study made by Harris there appears to be a decided relationship between ability to spell words of Latin derivation and the number of years devoted to Latin. In this case, however, it happened that Professor Irving King had previously given intelligence tests to these same students by means of which it is possible to secure a more accurate notion of the forces really producing the better scores of the Latin students. The following table shows the relation between the index of intelligence (that is, the percentage above or below the average adult intelligence), the number of years spent in the study of Latin and the ability to spell words of Latin derivation for the two extreme Latin groups.

TABLE 69. After Foster ('17)

Best 14 of 4	year	Latin	girls,	mental	ability	av.	25.8,	spelling	av.	5%	error
Poorest 13	"	"	"	"	"	"	−9.5,	"	"	33%	"
Best 12	"	"	boys	"	"	"	13 4,	"	"	7%	"
Poorest 12	"	"	"	"	"	"	−12 4	"	"	44%	"
Best 10 of o-	year		girls	"	"	"	19.1,	"	"	11%	"
Poorest 10	"		"	"	"	"	−14.2,	"	"	34%	"
Best 10	"		boys	"	"	"	5 0,	"	"	23%	"
Poorest 10	"		"	"	"	"	−27.2,	"	"	63%	"

This table shows clearly that students who chose to study Latin had on the average a distinctly better native intelligence than the non-Latin students and that the ability to spell words of Latin derivation was to a considerable extent due to this superior intelligence rather than to the study of Latin.

The secretary of the College Entrance Examination Board made an extensive tabulation of the records of the classical and the non-classical students who took the examinations in 1914, 1915, and 1916. The classical students are the ones who offered Latin or Greek, or both; non-classical students are those who offered neither Latin nor Greek. A total of 21,103 candidates are concerned in the following table which is based on the marks in all subjects except Latin and Greek (reported in *Value of Classics*, 1917, p. 366):

COMBINED RATINGS IN ALL THE NON-CLASSICAL SUBJECTS

Candidates who obtained a rating of 90 to 100:

2.95% of all the classical candidates.
2.05% of all the non-classical candidates.

The classical students show a superiority of 44%.
Candidates who obtained a rating of 75 to 89:

17.31% of all the classical candidates.
12.31% of all the non-classical candidates.

The classical students show a superiority of about 40%.
Candidates who obtained a rating of 60 to 100:

51.96% of all the classical candidates
40.97% of all the non-classical candidates.

The classical students show a superiority of about 27%.

A committee in connection with the Princeton Conference made a comparison of the honors received by classical and non-classical students upon graduation from high schools, academies and colleges. The table is based upon 2,799 classical and 5,606 non-classical students from 19 high schools and academies, and upon 4,092 classical and 2,003 non-classical students from 17 colleges and universities:

"The combined data from the nineteen high schools and academies reporting yield the following results:

"Students receiving High Honors at Graduation were 18% of all the classical students, but only 7.2% of all the non-classical students.

"That is: the classical students show a superiority of 150%.

"Students receiving Honors at Graduation were 32.1% of all the classical students, but only 30.8% of all the non-classical students.

"That is: the classical students show a superiority of 36.7%.

"Students receiving Honors or Prizes for Debating, Speaking or Essay-writing were 8.8% of all the classical students, but only 3.5% of all the non-classical students.

"That is: the classical students show a superiority of 150%."

"The combined data from the seventeen colleges and universities reporting yield the following results:

"Students receiving High Honors at Graduation were 17.3% of all the classical students, but only 6.6% of all the non-classical students.

"That is: the classical students show a superiority of 162%.

"Students receiving Honors at Graduation were 46.5% of all the classical students, but only 38.5% of all the non-classical students.

"That is: the classical students show a superiority of 20.7%.

"Students elected to Phi Beta Kappa were 16.8% of all the classical students, but only 8.9% of all the non-classical students.

"That is: the classical students show a superiority of 88.8%.

"Students winning Prizes or Honors for Scholarship in Other than Classical Subjects were 13.5% of all the classical students, but only 9.3% of all the non-classical students.

"That is: the classical students show a superiority of 45.2%.

"Students serving on the Editorial Boards of Student Newspapers and Magazines were 15.1% of all the classical students, but only 9.2% of all the non-classical students.

"That is: the classical students show a superiority of 64.1%.

"Students acting as Members of Intercollegiate Debating Teams were 5.1% of all the classical students, but only 3.2% of all the non-classical students.

"That is: the classical students show a superiority of 59.4%." (Reported in *Value of Classics*, pp. 381–383.)

These statistics are interesting enough, but they represent a mingling of training and native ability in the superiority shown and as such do not afford conclusive proof for the efficacy of classical training.

Wilcox ('17) made an inquiry with the endeavor to ascertain the amount of difference in original capacities. If the superiority of the Latin students is due to their study of Latin, we ought to find that they were not superior, or at least not as much superior, to the other students before they undertook Latin. If, however, we should find that the Latin students were as superior before they studied Latin as afterwards, we may infer that Latin had nothing to do with their superiority.

Wilcox tabulated the records of pupils in the Iowa City High School graduating during a period of ten years. He tabulated separately the grades made by all the students with Latin or German, but not with both. These results are shown in the following table in which the numbers were obtained by transposing the symbols E, G, M, P, and F into numerical values of 4, 3, 2, 1, and 0 respectively.

TABLE 70. After Wilcox ('17)

Median grades in English of Iowa City High School students who studied Latin or German

	Freshman	Sophomore	Junior	Senior
All Latin (184).........	6.21	6.29	5.80	6.11
All German (120).......	4.93	5.29	4.72	4.92
Difference....	1.28	1.00	1 08	1.19

The comparison is graphically shown in Figure 56. It will be noticed that the superiority of the classical group is found in the freshman year and continues throughout the course.

A comparison was also made of the English grades of students having four, three, and two years of Latin. This is shown in the following table:

TABLE 71. After Wilcox ('17)

Median grades in English of Iowa City High School students who had 4, 3, and 2 years of Latin

	Freshman	Sophomore	Junior	Senior
4 yrs. Latin (31)........	7.14	7.37	6.70	7.33
3 yrs. Latin (27)........	6.33	6.56	6.30	6.45
2 yrs. Latin (126).......	5.96	5.93	5.39	5.50

A graphical comparison is shown in Figure 57. It is evident that those who were destined to take four years of Latin were already in their freshman year clearly superior to those taking less Latin.

Wilcox made a similar investigation of the graduates of the high school of Cedar Rapids, Iowa.

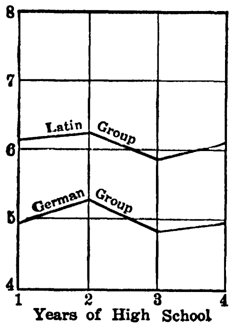

FIG. 56.—Median grades in English in the four successive high-school years of students taking Latin or German. After Wilcox ('17).

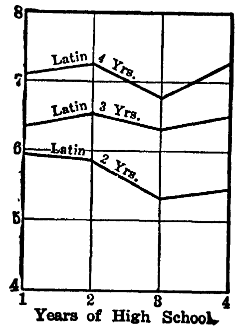

FIG. 57.—Median grade in English in the four successive high-school years of students who had 4, 3, or 2 years of Latin. After Wilcox ('17).

Comparisons were made of the grades of 150 graduates having Latin, German or no foreign language. This is shown in Table 72.

TABLE 72. After Wilcox ('17)

Median grades in English of Cedar Rapids High School students who studied Latin, German or no foreign language

	FRESHMAN	SOPHOMORE	JUNIOR	SENIOR
All Latin (70)	91	89	89	86
All German (60)	86	86	85	87
No foreign language (30)	86	83	84	85

It will be observed that in the freshman year the classical group was superior to the other two, but that by the senior year there was very little difference in the three groups.

In Table 73, comparison is made of the English grades of students having four, three or two years of Latin.

TABLE 73. After Wilcox ('17)

Median grades in English of Cedar Rapids High School students with 4, 3, or 2 years of Latin

	Freshman	Sophomore	Junior	Senior
4 years Latin (19).......	94	94	93	92
3 years Latin (9)........	87	88	87	85
2 years Latin (38).......	90	89	88	84

Here again the people with four years of Latin maintain their superiority throughout the course.

"Conclusions: It seems evident, so far as the Iowa City and Cedar Rapids high schools are concerned, that the frequently demonstrated superiority of students who have had Latin is not due to the special discipline or training secured in the study of Latin. It is probably due to the fact that, as a whole, the students who elect Latin are somewhat superior to those who refuse to take it."

Perkins ('14) made an investigation to determine the effect of emphasizing the derivation of English words from Latin words in the instruction of Latin in the commercial course in the Dorchester, Massachusetts, high school. This investigation was designed to eliminate as far as possible the differences in original abilities between pupils with Latin and without Latin. His report follows:

"Obviously, the first step was to select two sets of pupils of equal ability, one set in the second year of Latin, and the other in the second year of a modern language. Accordingly, we chose pupils such that each group had virtually the same average mark in Latin, on the one hand, and modern language, on the other, and also in English, with the result, in actual figures, that the non-Latin group in the two studies averaged 0.5 of 1% the higher. To make doubly sure that the Latin pupils were not favored, the non-Latin group were taken from the section of Mr. Murdock, a classical scholar, who in his English teaching emphasizes the Latin element in the language. There were twenty-one pupils in each set, all in the second year class of the school.

"Five measurements were made, one in spelling, one of the use of words in sentences, the third in definitions and parts of speech, the fourth in the meaning of words and spelling, and the fifth in excellence in vocabulary.

"Miss Humphrey selected the words in Nos. 1–4, and the subject in No. 5. In Nos. 1 and 2 the words were taken from the 600 or 800 derivatives in the notebooks of a fourth-year pupil of the class, who was excluded from the measurements. Moreover, to be fair to the non-Latin group, care was taken not to select words too difficult. In No. 3 the words were taken from the 'Tale of Two Cities' which the pupils of both groups were reading at the time in connection with their work in English II. Of the twenty words in No. 4, ten were taken from the 'Tale of Two Cities' and ten from other sources. The subject in No. 5 was, 'What I like to do best.' The papers were marked by teachers in the English department and the results given to me. Altogether, six teachers of English assisted in the measurements.

"To these five measurements is added a sixth—in my opinion most impressive of all. This test was made last June, shortly after I had received Professor Holmes's letter, by Miss Gormley, with her pupils in English II. As it happened, Miss Gormley, who was also the 'home-room' teacher of all the pupils and consequently had access to their marks, in making up the two groups to be composed of pupils of equal ability, took into account not only foreign language and English II, as was the case in measurements 1–5, but also the studies the pupils had taken during the year. Hence we have even more reason in this case than in the others to assume that the pupils were of equal ability. In each set there were seventeen second-year students. The words were taken entirely from Franklin's *Autobiography* and *Silas Marner* which all were reading at the time. The Latin pupils were selected from the first class I had had in the subject, just as they were completing the course at the end of the second year.

"The result of the six measurements were as follows:

	Averages	
	Latin Per Cent	Non-Latin Per Cent
January and February, 1914:		
1. Spelling	82.5	72.6
2. Use of words in sentences	57.5	40.6
3. Definitions and parts of speech	69.5	33.3
4. Meaning of words and spelling	57.0	27.5
5. Excellence in vocabulary	36.0	6.8
June, 1913:		
6. Meaning of words and spelling	63.3	12.3
	367.8	193.1
Averages	61.3	32.18
	32.18	
Difference	29.12%	

"In No. 1, the spelling measurement, the words were not difficult, but such as ordinary pupils of sixteen should know something about, whether they had studied Latin or not—as 'valedictory,' 'competition,' 'occurrence,' 'benevolence,' 'legible.'

"In No. 2, the pupils composed sentences containing the derivatives, some of which, in this measurement also, ought not to be unfamiliar to non-Latin pupils in their second year of English, as 'impediment,' 'advocate,' 'reference,' 'anticipate,' 'subside.'

"In the third measurement, the difference in the averages of the two groups—69.5% and 33.3%—was so great that Miss Humphrey thought that perhaps too difficult words had been selected, or at least words which placed the non-Latin students at an unreasonable disadvantage. Curiously enough, in this measurement the words were taken, not from the notebooks of a Latin pupil, as in the first two tests, in which the difference between the two groups was much less, but, as stated above, from *The Tale of Two Cities*. Furthermore, in No. 3, the non-Latin pupils were so far afield in giving accurate definitions, and so confused in classifying the words as to parts of speech, that it was decided to give another test in which they should be asked, not to define words, but to give their meanings, with the parts of speech omitted entirely. The results in this measurement—57% and 27.5%—were virtually the same as in No. 3.

"Since practically every second-year pupil could write at least passably on such a subject as 'What do I like to do best' it was decided to make the basis of comparison in No. 5, not the average of the two groups, but the percentage of rating above the passing mark. Moreover, in this vocabulary test, emphasis was laid, not merely upon words of Latin origin, but upon any words out of the ordinary, from whatever source. The wide difference in the results from the view-point of excellence in vocabulary—36.0% and 6.8%—shows clearly what I have always believed and maintained, namely, that the work in commercial Latin necessarily gives the pupils the dictionary habit, the results of which extend far beyond the Latin derivatives actually studied.

"Of all the measurements, No. 6, was perhaps the most convincing. In this test, the Latin pupils, unlike those in Nos. 1–5, had had during the last six months of the two years' course the benefit of drill in a vocabulary not in the commonest use and yet valuable and even necessary to educated people. The list of words was taken entirely from Franklin's *Autobiography* and *Silas Marner* which the pupils had just read, and was not of unusual difficulty, consisting of such words, for example, as asperity, promiscuous, mortuary. Yet by referring to the results it will be seen that to the non-Latin group of pupils such words were practically meaningless.

"An examination of the marks on these tests may prove of interest. Among the seventeen non-Latin students the highest grade was 30%, and five zeros were recorded. In the Latin group, on the other hand, the lowest mark was 30%, while one pupil received 100%, two 90%, two 80%, five 70%, and only three had below 50%. The difference in averages of the two groups was 53%." (Perkins, '14, pp. 11–14.)

This investigation is interesting and one of the few whose results were carefully worked out to make a precise comparison after eliminating differences in original capacity. It is a question, however, to what extent the Latin students may have been favored by the manner in which the words for the various tests were selected. For tests 1 and 2 they were taken from the lists compiled and studied by the Latin pupils. Even when the words are selected from English sources such as *Silas Marner* and Franklin's *Autobiography* there is still the question as to the particular words chosen for the test. It is obviously unfair to select words which are relatively rare and whose meaning may readily be inferred from their origin. To what extent the words were selected fairly cannot be judged since Perkins does not give the lists of words used.

The writer ('17) undertook a study to determine as precisely as possible, the relative shares contributed by language training and by original ability toward proficiency in English composition. A series of tests was carried out with a group of 177 university students. These tests together with their findings are given in Table 74.

No. 3 consisted in writing an extemporaneous composition within a limited time. These compositions were rated by three judges by the Hillegas Scale.

No. 4 gives the average number of words written by each group of students.

No. 5 gives the average number of different words used in each composition.

No. 6 was a test in speed of reading. The numbers refer to the words read per second.

No. 7 gives the number of words written in reproducing the thought of the passage read in No. 6.

No. 8 gives the number of A's canceled in one minute in the well-known A-test.

No. 9 gives the scores made in canceling in one minute a certain geometrical form on a page of similar forms.

No. 10 consisted in reading to the persons a series of ten words to see how many they could recall immediately afterwards.

No. 11 was carried out by giving a stimulus word and having the persons write as many associated words as they could in thirty seconds.

No. 12 consisted in giving a series of ten words and allowing fifteen seconds to each word for writing as many synonyms as possible.

No. 13 was a set of tests in imaging geometrical forms.

No. 14 consisted in presenting to the subjects seven words, one at a time, spelled orally backwards by the experimenter. The persons wrote down the words which they could thus recognize.

TABLE 74. After Starch ('17)

						Percentage)–15 Group 1–2 Group		
1. Years of foreign language.........	1–2	3–4	5–6	7–8	9–15		No Latin	Latin
2. Number of persons .	14	53	49	40	21	....	59	112
3. Composition (Hillegas scale)	67.6	69.3	68.7	71.5	78.2	15.7	67.9	72.6
4. Words written.....	140.7	158.3	162.1	168.1	181.4	28.9	158.7	165.0
5. Different words used	82.3	86.6	96.6	96.8	111.6	35.4	89.6	95.2
6. Reading speed	4.5	5.4	5.5	5.2	6.0	33.3	5.4	5.5
7. Reading comprehension.............	60.0	65.0	70.5	65.7	68.2	13.7	64.0	69.0
8. Perception A-test..	66.2	66.8	66.2	67.8	66.1	0.0	66.2	67.1
9. Perception form ...	7.5	7.7	7.8	7.5	8.0	6.7	7.8	7.9
10. Memory words.....	7.4	7.2	7.3	7.2	7.2	−2.7	7.3	7.3
11. Association free....	23.5	25.9	22.6	29.9	28.7	26.4	25.3	26.3
12. Association synonyms........	15.4	15.0	15.9	15.4	14.2	−7.7	15.0	15.4
13. Imagery forms.....	7.0	7.6	7.0	7.1	7.8	11.3	7.2	7.4
14. Imagery words....	5.0	5.7	5.1	5.7	6.1	21.4	5.5	5.6
15. Years of English..	5.1	4.9	5.3	5.5	5.5	....	5.0	5.3
16. Grades, first year of high school.......	83.0	85.7	83.7	86.7	88.0	....	84.5	85.7

A general inspection of the table shows that there is a steady increase in the scores of practically every test from left to right with increasing years of language study. Thus the 9–15 year group wrote considerably better and longer compositions than the 1–2 year group. The purpose of tests Nos. 6–14 was to ascertain to what extent this superiority was due to original superiority of ability or to the effect of language training. Tests 8–14 were selected particularly because the capacities to do them would probably be affected very slightly if at all by language training. These show on the whole a distinct superiority of inherent ability on the part of the groups who studied languages for longer periods of years.

In order to make a crucial comparison as to how much of the greater composition ability was due to the greater original capacity of the pupils and how much was due to their greater training in language, the grades received by these students in all the subjects carried during the first year of the high school were obtained from the entrance records of the University. The amounts of difference in original ability of the groups who later pursued varying amounts of language work would be definitely indicated by this method, since at that time none had had more than one year of foreign language. The difference in the scholastic grades at the end of the first year of the high school between those who later pursued languages for a total of 9–15 years and those who pursued languages for a total of 1–2 years could certainly not be due to language training.

Row 16 gives for the different groups the average scholarship grades during the first year of the high school. It will be noted that there is a steady increase from group to group. The 9–15 year group had an average grade of 88.0, or five points higher than the 1–2 year group.

The next problem was to compare in common terms the five points of difference in scholarship on the percentage scale with the difference of 10.6 in quality of composition as measured by the Hillegas scale. To reduce these two types of measurements to commensurate units, fifty-eight compositions were rated by four persons both by the percentage method and by the Hillegas scale. By a process of equating values it was found that 1.0 point on the percentage scale equals 2.1 on the Hillegas scale. The difference of five points, percentage scale, in original scholarship between the 1–2 year group and the 9–15 year group would be 10.5 in terms of the Hillegas scale. The surprising result seems to be that the difference of 10.6, Hillegas scale, in quality of composition between these two groups is approximately equaled by 10.5, the difference in original scholarship when expressed in terms of the Hillegas scale. The conclusion seems, therefore, unavoidable that the difference in ability in English composition is due practically entirely to a difference in original ability and only to a slight or no extent to the training in foreign languages. [For the method of this computation the reader is referred to the original report of this investigation. ('17.)]

The increase in length of composition and in speed of reading is large and very probably in excess of the difference in original ability. Training in foreign language seems to have produced a distinct effect in greater fluency of words in writing and in more rapid perception of words in reading.

Miss M. Theresa Dallam ('17), a teacher of English in the Western High School of Baltimore, made a series of experiments on her pupils to test the truth of her conviction "that Latin classes do much better work in English than the classes that have not studied Latin." Out of 114 students she selected 34 fourth-year students, 17 Latin and 17 no-Latin or modern language pupils, by pairing them on the basis of their general scholarship records, so that the two groups would be equal in mental ability. The Latin group had an average of 78 and the no-Latin group 79. Kelley's Silent Reading test was also made with them. The Latin group made an average score of 31.1 and the no-Latin group 33.1. The two groups were, therefore, very nearly equal in general ability. The Latin group had studied the language for four years.

Miss Dallam then made seven tests: spelling, reproduction, dictation, Latin derivations, definitions, compositions, and English grammar. The results were as follows:

TABLE 75. After Dallam ('17)

Average percentages attained

	Sp.	Repro.	Dict.	Deriv.	Defin.	Comp.	Gram.
Mod. Lan. Group.........	89.2	63.0	96.0	29.3	73.0	71.8	59.7
Latin Group.............	90.7	65.9	95.0	52.0	75.2	75.4	67.6

Thus the Latin group made a distinctly higher average in derivations and grammar, an appreciably higher grade in compositions; a slightly higher average in spelling, reproduction, and definitions; and a slightly lower average in dictation.

Miss Dallam then computed coefficients of correlation for the Latin group between their grades in Latin and each of the seven tests, and for the modern language group between their grades in modern languages and each of the seven tests. These correlations were as follows:

TABLE 76. After Dallam ('17)

Coefficients of correlation

	Sp.	Repro.	Dict.	Deriv.	Defin.	Comp.	Gram.
Mod. Lan. Group..	+.09	+.19	−.04	−.02	+.23	+.11	+.28
Latin Group......	+.05	+.15	−.16	+.13	+.15	+.28	+.46

These correlations are so low that, with the exception of the ones for grammar, no significance can be attached to them except to say that there is practically no correlation between the various comparisons made and that the Latin group shows no superiority over the modern language group in spelling, reproduction, definitions, and a doubtful superiority in dictation, derivations, compositions, and grammar. The differences that are shown are non-committal and so small that they would have to be substantiated with other groups to be conclusive.

d. Science. Claims of general training to be derived from the pursuit of the sciences are practically as extensive and confident as those made for the languages. The only difference is that the claims for the sciences have not been questioned or investigated as much as those made for the languages. Bagley has summarized them in the following manner:

"1. The formation of some useful specific habits,—through training, routine, rationalized practice.

"2. The acquisition of useful information,—through methodical study, instruction, and drill.

"3. The adoption of valuable ideals, or 'emotionalized standards,'—inculcated through the inspiration to be gained from the teacher, from the lives of great scientists, and from experiences of intimate contact with nature.

"4. The acquisition of facility in the use of facts, ideas, and methodical thought processes, for the solution of problems, the overcoming of difficulties, and the accomplishment of worthy purposes,—through the mental discipline afforded by properly graded practice in the solving of scientific problems.

"5. The development of taste, and power of appreciation,—to be gained through a clear apprehension of unity, adaptation, economy, order, and system in nature as interpreted by science.

"6. The development of scientific or philosophic insights, perspectives, and attitudes of mind that serve as safeguards to the intelligent interpretation of contemporary life,—through acquaintance with systems of organized knowledge." (Bagley, '11.)

One of the important values claimed for the sciences is the general training of accuracy and fidelity in observation and the transference of this particular type of training to other types of observation. Miss Hewins ('16) made an attempt to measure the extent to which this improvement is general or carries over to other types of observation. She divided each of three classes in botany, composed of 34 boys and 50 girls, in the first year of a New York high school, into two groups, and gave them a series of tests in various kinds of observation of a biological and non-biological nature as follows:

Series 1

No.	Date	Test Series	Exposure	Time for Recording
1	April 22	Horse chestnut stem		10 minutes
2	23	Picture	30 seconds	5 "
3	23	Forsythia flower		5 "
4	24	10 syllables	30 "	1 "
5	24	Lilac leaves		8 "
6	25	Nonsense figure	30 "	1 "
7	25	Geometrical figure	30 "	3 "
8	26	10 2-place figures	30 "	1 "
9	29	Scouring rush	1 minute	5 "
10	29	Maple seedling	1 minute	5 "
11	30	Pea chart	30 seconds	5 "
12	30	Figure in air		
13	May 1	Potato chart	30 seconds	5 "

Practice series

		Material	Recording
1	May 15	Description of lilac flower	
2	16	" " box-elder leaf	10 minutes
3	17	" " the stem, leaf, and flower of gill-run-over-the-ground	for each test
4	20	Description of flower stalk and flowers of the lily-of-the-valley	
5	21	Description of the horse chestnut flower	
6	22	" " " buttercup flower	
7	23	" " " mustard flower	
8	24	" " " dogwood flower	
9	27	" " " deutzia flower	
10	28	" " " columbine flower	

Series 2

		Test Series	Exposure	Time for Recording
1	June 3	Sassafras stem		10 minutes
2	4	Picture	30 seconds	5 "
3	4	Syringa flower		5 "
4	5	10 syllables	30 "	1 "
5	5	Forsythia leaves		8 "
6	6	Nonsense figure	30 "	1 "
7	6	Geometrical figures	30 "	3 "
8	7	10 2-place figures	30 "	1 "
9	10	Moss plant	1 minute	5 "
10	10	Pumpkin seed chart	1 minute	5 "
11	11	Grape chart	30 seconds	5 "
12	11	Figure in air		
13	12	Wild carrot	30 "	5 "

The practice series was continued for a period of ten days with one section of each of the three classes. The other section of each class answered questions from books on the material of the lessons. At the end of the period, she gave the tests in Series 2 to both sections of each class and obtained the following gains in scores over Series 1:

TABLE 77. After Hewins ('16)

Gross gains in scores

	Biological Tests	Non-biological Tests
Practiced groups:		
Boys	8.06	8.97
Girls	6.41	6.20
Average	7.23	7.58
Unpracticed groups:		
Boys	3.03	5.37
Girls	−1.24	5.60
Average	.89	5.48
Residual difference between practiced and unpracticed groups	6.34	2.10
Percentage gains of residual differences over scores in first test series	33.9%	5.4%

(Median score in biological tests in Series 1 before practice 18.7)
(" " " non- " " " " " " " 39.0)

Miss Hewins infers a considerable gain in both the biological and non-biological types of observation.

"Table 77 shows that in the biological tests, the average gain of each practised boy was 8.06 per test for the 5 tests while the unpractised showed a gain of 3.03. The practised girls averaged 6.41 gain per test, while the unpractised lost 1.24 per test. In the non-biological, the practised girls gained 6.2 per test for the 6 tests while the unpractised gained 5.6; the practised boys gained 8.97 per test and the unpractised boys gained 5.27 per test."

"Feeling that the balance of arguments and scientific proofs were against formal discipline when this investigation was begun, I am forced by the results obtained to admit that in this experiment the proof seems to be on the affirmative side.

"A valuable lesson, I think, can be drawn from one phase of this investigation. By consulting the tables and summaries, it will be seen that sometimes one division does not fall in line with the general trend, but that a larger number outweighs the negative and shows positive results." (Hewins, pp. 111 and 113.)

In order to make a relative comparison of these gains on the basis of the original scores in the first series of end tests, I have computed the percentages of gain in these end tests as shown at the bottom of the above table. When such a relative comparison is made, the net gain in the non-biological observations is very small, being only 5.4% as compared with 33.9% in the biological observations, or about one-sixth as much in the non-biological as in the biological types of observation. The improvement in the biological end tests can hardly be counted as evidence of spread of training since these tests were so similar to the training series that they were all but identical with it. The spread of training is apparently not as great as Miss Hewins believes it to be.

e. Geometry. Geometry, particularly in the high school, is urged to possess a considerable amount of disciplinary value:

Plato (Republic, Book 7) emphasized this opinion thus:

"Moreover, the science (Geometry) has indirect effects which are not small.

"'Of what kind?' he said.

"'There are the military advantages of which you spoke,' I said; and in all departments of knowledge, as experience proves, any one who has studied geometry is infinitely quicker of apprehension than one who has not.

"'Yes, indeed,' he said, 'there is an infinite difference between them."

Rugg ('16) made a study of the spread of training in the case of learning descriptive geometry. He performed three groups of tests at the beginning and at the end of the course in descriptive geometry, in February and in June respectively, with 326 students in the College of Engineering of the University of Illinois. The same tests were made with 78 students in other colleges as a control group who did not pursue the course in descriptive geometry. The three groups of tests were of a non-geometrical, Tests 1 and 2, quasi-geometrical, Test 3, and strictly geometrical nature, Tests 4 and 5. Partial illustrations of the various tests follow:

Test 1.

1. Divide eighty-one by seven.
2. Divide seventy-eight by four, etc.

Test 2.

1. Divide eight sixty-two by three.
2. Divide seven ninety-five by four, etc.

Test 3.

A test in imaging letters.

Test 4.

The Painted Cube Test.

A three-inch cube, painted on all sides, is cut into one-inch cubes.

1. How many one-inch cubes have paint on three sides?
2. " " " " " " " " two " ?
3. " " " " " " " " one side?
4. " " " " " " " " no side?

Test 5.

Geometrical Objects' Test.

The problem: Form a mental picture of each object and count the number of straight lines which it would take to construct each one in space.

1. A wedge.
2. Four triangles attached to a square, bases coinciding with the sides of the square, etc.

The main results are represented in the following table:

TABLE 78. Representing the results for "Rights." Adapted from Rugg's table, p. 123 ('16)

	Test 1		Test 2		Test 3		Test 4		Test 5	
	Training Group	Control Group	Training Group	Control Group	Training Group	Control Group	Training Group	Control Group	Training Group	Control Group
Feb. scores	20.0	17 38	20.00	19.3	4.67	4 50	0 59	0.28	2.32	1.92
June scores	22.81	20 06	23.00	19.15	5.66	4.58	1.50	0.68	3 33	1.81
Gross gain	2.81	2.68	3.00	−.15	.99	0.083	.91	0.40	1.01	−0.10
% gain	14.5	15.6	15.0	−.78	21.2	1.8	157.0	143.0	43.5	−5.0
Residual difference	−1.10		15.78		19.4		14.0		48.5	
Averages	7.34				19.4		31.7			

Rugg concludes that the residual gain indicates a considerable transfer of improvement to all three types of abilities tested.

"The study of descriptive geometry (under ordinary class room conditions throughout a semester of 15 weeks) in which such natural and not undue consideration is given to practice in geometrical visualization as is necessary for the solution of descriptive geometry problems operates:

"(1) Substantially to increase the students' ability in solving problems requiring the mental manipulation of a geometrical nature, the content of which is distinctly different from the visual content of descriptive geometry itself.

"(2) Substantially to increase the students' ability in solving problems requiring the mental manipulation of spatial elements of a slightly geometrical character, i. e., problems utilizing the fundamental elements of geometry (the point, line, and plane), but apart from a geometrical setting, and in such form as to offer no geometrical aids in solution.

"(3) Substantially to increase the students' ability in solving problems requiring the mental manipulation of spatial elements of a completely non-geometrical nature, i. e., problems in which the straight line and plane do not appear in any way whatever.

"(4) The training effect of such study in descriptive geometry operates more efficiently in those problems whose visual content more closely resembles that of the training course itself, i. e., in those problems whose imagery content is composed of combination of points, lines, and planes, and in which the continuity of the manipulating movements approaches the continuity of those in the training course. (Rugg, pp. 114–115.)

These results are not very different from those surveyed in the preceding chapter. Rugg states the results perhaps as favorably as the data permit, perhaps too favorably. At any rate we ought to note that the gain in the non-geometrical tests, Nos. 1 and 2, is only about one-fourth or one-fifth as much as in the strictly geometrical tests, Nos. 4 and 5. We also should note, as Rugg himself points out, that only about two-thirds of the persons in the training group gained; the remaining one-third did not gain or show transfer.

General Interpretation. The transference of training of the capacities involved in the learning of school material is very small so far as present partial data indicate. This seems to be equally true of all school subjects, the sciences as well as the languages and mathematics. If we represent the possible transfer effect as ranging from 0%, or none, to 100%, or an improvement in other capacities equal to that in the capacity trained, then the amount of transfer is much nearer to the 0% end than to the 100% end. It probably is 0 or very nearly 0, for all capacities which are not distinctly similar or related to the capacities specifically trained. Thus in the author's experiment, practice in mental multiplication improved other forms of mental calculations about one-fourth as much, but had no effect upon immediate memory of numbers or words. In Winch's experiment, practice in arithmetical computations had either no effect or a doubtful effect upon arithmetical reasoning. In Rugg's investigation, practice in visualization ordinarily done in a course in descriptive geometry had only a moderate effect upon visualization of other sorts. In Miss Hewin's study, improvement in biological observation improved non-biological observation only one-sixth as much. In the author's investigation, the study of foreign languages had no effect upon the capacity to write an English composition, the study of Latin seemed to produce a small increase in English vocabulary and a decided increase in the knowledge of English grammar, but only a very small increase in discrimination in correct English. Wilcox likewise found that the study of Latin had no appreciable effect upon work in English classes as measured by marks. The superiority of the Latin pupils was due to their superior native ability rather than to the study of the language. Perkins found that Latin as taught by him with special emphasis upon word derivations and meanings produced a noticeable increase in the ability to spell, define, and use English words.

The transfer effects of the training of the abilities in school subjects is very much less than is commonly assumed. This is probably due, in the first place, to the fact that the improvement in the capacities exercised specifically by the school subjects is usually not as great as is commonly believed by teachers. The modifications produced in the minds of the pupils are considerably less than teachers usually assume as judged by the modifications produced in their own minds after much longer and harder study. To illustrate, teachers are inclined to believe that a course in mathematics has produced a much greater improvement in mathematical reasoning, or that a course in history has brought about much greater facility in handling and interpreting historical material, or that a course in psychology has brought about much keener insight into the operations of mental process, than these respective subjects actually have produced. This is probably the result of the teacher's naturally egotistical belief in the effectiveness of his own work.

The small effects of transfer are probably due, in the second place, to the fact that the conditions for securing transfer are not as favorable on the whole in the case of school subjects as in the case of the special laboratory experiments on transference.

The evidence on spread of training in school material tends to support for the most part the theory of identical elements. The effects are the largest where there is similarity or identity of material as, for example, in the case of the effect of the study of Latin upon the study of Spanish, or upon the knowledge of English grammar. The fact of identity of material or similarity of procedure makes possible a greater control of the spread of improvement through methods of teaching whereby the identity or the use of identical material may be emphasized in as many desirable relations as possible. This is illustrated by the spread of the effects of Latin as taught by Mr. Perkins.

In formulating an opinion concerning general training effects resulting from training of special capacities, we must bear in mind that even where the transfer effect is considerable, as much as one-fourth to one-third as much as in the capacity specially trained, it is obviously more economical to give practice directly to the capacities which we want to train rather than to do it indirectly with the hope that the improvement may be transferred to them. Concretely, even if the study of Latin under favorable methods of teaching does improve the spelling of English words, would it not be more economical to study directly the spelling of the words which are to be acquired? Knowledge of the most common Latin words from which the largest number of English words are derived could be obtained in a relatively short period of time, probably a year or even less. Learning to play the piano might help in learning to play the violin, but no sane person would devote very much time to the piano if his sole purpose is to learn to play the violin.

Even if mathematics may cause some improvement in reasoning about bargains, even if the study of Latin may increase English vocabulary, or even if a study of animal psychology did make a man a better teamster, these effects are relatively very small and can be produced much more economically by a direct study of bargains, or of the origin and meaning of English words, or of driving horses. A course in mathematics or in Latin or in psychology will have to stand primarily on its own feet for the content that it offers or the skill that it develops. These by-products may be useful but they cannot be the sole purpose of the efforts put into a course. The value of a meal depends upon the meal itself and not upon the crumbs that fall from the table. Whenever a subject loses its content value through changed social conditions it seems mysteriously to acquire a great deal of disciplinary value.

An immense amount of confusion in the thinking about the problem of mental discipline and the value of school subjects, even on the part of distinguished thinkers, has resulted from a failure to discriminate between the effect of a certain kind of education and the native capacities of the individuals subjected to the education. Whenever allowance or deductions for differences in original ability have been made, the general disciplinary effect has been found to be much less, or, in many instances, even non-existent. To argue that because certain great leaders of men had a certain type of education, it must have produced their greatness does not prove the point. They probably would have achieved distinction if they had had any other sort of education. If the chief argument for pursuing a given subject is that it selects the more able pupils, it would be much more economical to do so by a shorter and more certain method. Almost any fifteen or twenty mental tests that can be applied in a psychological laboratory in two hours would separate much more accurately the gifted from the stupid.

Finally, the upshot of the experimental and statistical inquiries into the transference of training is that effects of training are transferred in smaller amounts and within much narrower limits than has commonly been assumed. This does not mean that there is no general mental discipline in any form of training, nor that the doctrine of formal discipline has been "exploded" but rather that the actual limits of general discipline have been more accurately defined. These limits, to be sure, seem to be much narrower than many are inclined to believe. So far as the value of school subjects is concerned, it means that the content value of a subject must be the prime reason and the general disciplinary value the secondary reason for pursuing it.

Before leaving this discussion, two further points should be borne in mind: The first is that any effect of transfer, even if very slight, would probably be worth while if it extended to all or to a large number of capacities. If training in botanical observation improved all forms of observation in life ever so little, it might still be the best form of training in observation. But the implication of the evidence thus far at hand is that the spread seems to extend only to rather narrow limits. The second point is, that while the trend of the arguments here presented would be to reduce the time devoted to some subjects, particularly in high school and college, we must be sure that we put something better in their places. The advantage of some of the subjects that would suffer reduction is that they are well organized for teaching purposes. Some of the new substitutes are not well organized and offer neither form nor content. Transitions should be made gradually so that the new branches may become organized and extended, and the teachers properly trained.

PART III

THE PSYCHOLOGY OF LEARNING: B. OF SCHOOL SUBJECTS

CHAPTER XV

THE PSYCHOLOGY OF LEARNING SCHOOL SUBJECTS

Psychology and Teaching. If education consists in making changes in human beings, if psychology is the scientific study of the mental processes of human beings, and if teaching consists in the facilitation of the changes to be made by the school, then it is obvious that knowledge of how to bring about these changes in the most economical manner must be based upon an exact knowledge of the processes involved in these changes. A reliable pedagogy can be based only upon a reliable psychology of the processes concerned. Engineering did not become a science until the physiical and chemical knowledge of the processes involved in a given project were thoroughly understood. To build a Brooklyn bridge involves a precise knowledge of the laws of gravitation, the strength of materials, the means of supporting weights against the force of gravity, and the like. To compound an electric cell involves a precise knowledge of the chemical action of certain elements. To know how to destroy bacteria harmful to plant and animal life, it is necessary to understand the biological processes of the particular plant and animal life concerned.

The foundation science for sane and dependable methods in education is psychology in its broad sense. The great difficulty in establishing a reliable pedagogy is the fact that sure and detailed knowledge of the psychological processes and laws in learning the material of the school subjects is largely unknown. In the past the schools have proceeded largely by guess. The future will have to map out in detail the psychological steps involved in each school subject, and to submit these processes to direct experimental investigation. A science of engineering was impossible until it was discovered how the physical and chemical laws operated in the particular conditions under which the bridge had to be built or the cell had to be compounded. Knowledge of the law of gravitation was practically useless until it was discovered how it operated under given concrete conditions of materials, distances, and circumstances. The psychological laws of learning will be practically useless until we shall know how they operate under the concrete conditions of the school and with the materials to be learned in the school. The reaction times of numerous psychological processes have been studied in great detail for many years, but this knowledge is practically useless in giving to the educator scientific information by which he may proceed to facilitate the reactions of a child in learning to write or to read. This knowledge is useful in furnishing a general background, in pointing the way, and in supplying a general experimental technique. But even these must often be materially modified and adapted to the solution of a particular problem in a given field. The old-time pedagogy has fallen into disrepute because it has been almost wholly a matter of personal guesswork. The slate must be wiped clean and only those principles and laws whose truth has been fully proved can be recorded thereon.

Problems. If we grant that the method and procedure of the school should be based upon a sound psychology of the processes involved in learning the special materials of the school subjects, it follows that the fundamental tasks to be done are these:

(1) A thorough and complete analysis of all psychological processes involved in the learning of a given subject, or in the acquisition of skill in it, and of the order and manner in which these processes intermesh.

(2) The devising of means by which these processes may be measured and tested so that the facility in their operation may be determined quantitively.

(3) The discovery of the most economical procedures by which each particular step in the entire process may be developed.

If we wish to determine how to memorize a poem in the most economical and most permanent manner, it is important to know the perception, association, and reaction processes involved, the means for definitely measuring facility in these operations and the means of controlling these processes most efficiently. Teaching is mental engineering; it consists in managing the mental processes concerned in learning the materials and in acquiring the skill of the school in the most effective and most profitable manner.

In considering the psychology and pedagogy of school subjects in the succeeding chapters, this three-fold division of the problems will be made for each subject and the available information of each one surveyed so far as our present knowledge warrants.

CHAPTER XVI

READING

Processes or Steps Involved in Reading

If we trace, for analytical purposes, the successive steps from the external presentation of the visual stimuli of the printed words on through the complicated elaborations within the mind and back to the external expression of reactions in the pronunciation of the words, we can discern the following order or combination of elements:

(1) Reception upon the retina of the stimuli from the printed page.
(2) The range of the field of distinct vision on the retina.
(3) The range of attention in apprehending visual stimuli.
(4) The movements of the eyes.
(5) The transmission of the visual impressions from the retina to the visual centers of the brain.
(6) The establishment or arousal of association processes whereby the incoming impulses are interpreted.
(7) The transmission of the impulses from the visual centers to the motor speech centers.
(8) The transmission of motor impulses from the speech centers to the muscles of the vocal chords, tongue, lips, and related parts.
(9) Execution of the movements of the speech organs in speaking the words.

These are the steps as they occur in the developed reading process. It is obvious that they do not follow each other in a temporal order but that some, as for example (1) to (4), occur simultaneously. During the early stage of learning to read there occurs, simultaneously with steps (1) to (5), a parallel series of steps derived through the ear by which the child learns the pronunciation of the word, thus: (1) reception in the ear of auditory stimuli from the pronunciation of the word by the teacher, (2) transmission of the auditory impulses from the ear to the auditory center in the brain, (3) transmission of impulses between the auditory and the visual centers whereby the sound and the sight of the word become associated. Silent reading involves only the first six steps except in so far as incipient speech movements accompany it, in which case the remaining steps enter in part.

Such an analysis as this may seem detailed to an unprofitable extent; but, in fact, it might be made even more detailed, depending upon the extent to which we are able to discern and describe the minuteness of the neural and mental functions involved. The more complete and accurate our analysis and description of the steps is, the more sure our knowledge for managing these processes will be; and ultimately, that is what teaching amounts to: The efficient management of the psychophysical processes concerned.

The next problem is, How does each of the elements in the reading process operate? The truth is that concerning many of them we know at the present time very little or nothing with certainty or completeness. Concerning some of them, however, considerable definite knowledge has been accumulated in recent years. What the differences between an efficient and an inefficient reader are, or what the difficulties in learning to read are at each of the steps can be inferred partly, but only partly, from our present knowledge about these factors. It is, however, certain that the differences and difficulties are to be found in these and possibly additional processes. We shall examine each of these steps in turn and survey what definite knowledge is available.

(1) The reception upon the retina of the visual stimuli from the printed page depends obviously on the one hand, upon the size and kind of type, length of lines, indentation of lines, paper, and illumination; and, on the other hand, upon the inertia of the retina in receiving stimuli. We know, for example, that for adults and for children above 10 or 12 years of age, type smaller than 8 or 10 points is probably too small to be perceived easily. Likewise, type larger than approximately 10 points, spreads out upon too large an area of the retina to be perceived quickly in as large groups as possible. Experiments by Dearborn ('06) and others have shown that probably the most advantageous length of line is in the neighborhood of 2½ or 3½ inches, and that it is better to have the lines on a page uniform in length instead of varying in length as is often the case in reading-texts in which illustrations are set into the margins and the lines made to vary according to

the space around them. We do not know what the most advantageous size of type is for younger children who are beginning to learn to read. We feel that it ought to be larger than for older children or for adults, but we do not know definitely how much larger.

So far as the inertia of the retina to the reception of the stimuli goes, we may infer that it varies in different individuals probably according to the normal distribution and that it may be less in rapid than in slow readers, but no definite measurements have been made to ascertain the truth about it.

In an investigation, as yet unpublished, by C. L. Hull and W. R. Ames, an effort was made to determine the relative effect upon the eye of reading from papers of various colors, of various amounts of gloss and of various degrees of perfection of the inked impressions. Four kinds of paper were compared: a matte white paper, a pink newsprint paper, a blue newsprint paper, and a very glossy white paper. As measured by the Ingersoll glarimeter, these papers had the following glare or gloss values respectively, 18.5%, 41.5%, 42%, 73.5%. A fair sized book was printed, uniformly upon each kind of paper in such a way that the successive pages followed one another on a single band of paper. These were placed upon special reels in such a way that the pages would be at a uniform distance from the eyes, at a uniform angle and would have uniform illumination.

Three measures of the changes in the eyes produced by reading were taken:

1. The number of lines read during a 15-minute period.

2. The number of spontaneous winks per minute while reading. These were recorded by a special device unsuspected by the subjects.

3. Extent of failure to recover during a 15-minute reading period from artificially produced diminution in the distance at which the subjects were able to see faint parallel lines as distinct lines. This distance was determined before and after reading by an elaborate recording device of considerable precision. Four subjects were used.

The final averages for each kind of paper by each of the tests are shown in the following table. In addition, the relative rank of the four papers is given for each of the three measures on a scale of 10. Lastly these ranks are averaged for a final score of all three measures.

TABLE 79. After Hull and Ames. From a Thesis in the Library of the University of Wisconsin, 1917

	Final Averages			Final Ranks			
	Loss of Visual Acuity in Inches	No. of Winks per Minute	No. Lines Read per 15-Min. Period	Visual Acuity	Winks	Lines Read	Average Rank
Matte White..	1.8	7.5	415	1	3	3.6	2.5
Pink.........	3.0	7.2	422	10	1	1.	4.
Blue...... .	2.7	8.1	414	5.4	7 2	4.	5.5
Gloss White...	2.0	8.5	398	3.4	10.	10.	7.8

Despite rather striking differences between the results obtained by the third measure and by the other two, the results as a whole, as indicated by the final average rank, show that glare is the decisive factor in diminishing ocular efficiency. The two white papers are respectively the best and the worst of the set. The two colored papers which are intermediate in glare are also found intermediate in ocular efficiency. These results suggest that color as such has little or no influence one way or the other.

A microscopic examination of the texture and perfection of the inked impressions of the various papers revealed the glossy white papers to be the most perfect, with the matte white, the blue and the pink in decreasing order. This order indicates that within ordinary limits the perfection of the printed impression is of little consequence.

(2) Concerning the size of the distinct field of vision, Ruediger ('07) made some experiments with the tachistoscope. Speed of reading may depend upon the horizontal area of distinct vision in the retina. Ruediger, however, concluded that little correlation exists between the areas of color zones on the retina and visual acuity and other qualities of sight, that the field of distinct vision is reduced as much as one-half when the eyes are tired from reading and that there is only a slight correlation between visual acuity and the size of the field of acute vision.

Judd, McAllister, and Steele ('05) marked the eye near the pupil with a flake of Chinese white and then directly photographed the movements of the eye. They concluded that there is no correlation between the size of the horizontal field of acute vision and the rate of reading or the number of pauses in reading. The correlations found were —.06 and —.10.

(3) The range of attention in apprehending words and letters. Dearborn, in his photographic records of the movements and pauses of the eyes, found that the eye takes in, on an average obtained from five subjects, 1.64 words at one time, or at each fixation. He believes that the word is the unit of reading. This is corroborated by the fact that slight misspellings in words are often not noticed and that such words as "psychology" and "physiology" may easily be confused with each other because they are perceived as wholes. In many instances he found that long words take no longer time to be perceived than short ones do. The habit of grouping is apparently very important. "It is not the short words as such but the words which cannot be easily grouped with others, which necessitate separate fixations." He also found that the range of attention with slow readers is often only a syllable while with the fast reader it is words and phrases. There is thus a large and important difference which is probably highly significant in the development of rapid reading as will be pointed out later.

By presenting ordinary printed material for very brief intervals, Gray showed that the number of words which could be comprehended at a single exposure was much larger than the number actually read at a single fixation—often twice as large. This means that the areas of visual apprehension overlap considerably at successive fixations. This suggests that the number of fixations might be reduced considerably by practice. Gray tried this out experimentally with two children, one a good and the other a poor reader, and found it to be true. The poor reader decreased the number of pauses per line from 15.5 to 6.1 by a 20-minute practice period each day for 20 days.

In this connection the possibility of improving the range of visual apprehension becomes important. Some years ago Miss Aiken claimed to have increased enormously the range of visual apprehension of the girls in her school by special training. G. Stanley Hall wrote of it, "I would not have thought such rapidity and accuracy possible if I had not seen it." Whipple, in an attempt to verify these claims under laboratory conditions with adult subjects, failed almost completely. Gray, attempting by the same methods to enlarge the range of visual impressions of two sixth-grade children for printed words, also failed completely. He repeated the experiment, however, with two fourth graders and succeeded in approximately doubling it. He concludes that such training, to be effective, should take place not later than the fourth grade.

(4) The movements of the eyes in reading. Closely related to the scope of apprehension is the rapidity and the precision of the movement of the eyes. More experimental work has been done on these two elements in reading than on any others. It has been known for a long time that the eyes in reading or in examining any object do not move along smoothly and continuously but that they move by jerks and pauses apparently in a very irregular manner. The eyes take rapid glimpses or snap shots of successive portions of a line of print and then piece them together in obtaining the meaning. The reason for it is that the eyes cannot see objects distinctly while they are in motion and consequently they perceive little or nothing during that time. A person cannot see the movement of his own eyes in a mirror. As soon as one can see them distinctly, they are at rest again. Dearborn found that letters fuse when passed in front of the eyes at the rate at which they themselves move across a page. Dodge ('00) also found by experiments that, while the eyes were in motion, no sensation resulted even though the stimulus strikes the eyes.

A fairly good conception of the nature of eye movements, pauses, and fixations in reading may be obtained by taking an open book, placing a mirror on one page while some one reads the opposite page, and then observing the action of the eyes as reflected in the mirror. Javal ('79) counted the pauses of the eyes by means of a sound attachment to the eyelids. Laudolt ('91) counted them by direct observation. Erdmann and Dodge ('08) counted them by observing the eyes of their subjects in a mirror. They found more pauses with difficult than with easy reading material, and also more pauses in reading a foreign language than in reading one's native language.

The first successful attempt to record eye movements was made by Huey ('98–'00). He attached a plaster of Paris cup to the cornea which was connected with a light aluminum pointer. This, in turn, rested on the smoked drum of a kymograph on which the movements were registered. In this manner he attempted to study the nature and rapidity of eye movements and the nature and length of successive pauses. Dodge then developed a falling plate camera which photographed a beam of light reflected from the cornea of the eye. This method was also used by Dearborn. It has the important advantage over Huey's method in that it eliminates the attachment of anything to the eyes themselves, which quite likely interferes with the natural movements of the eyes.

Readings by five subjects of the same newspaper passage

1

Subjects		Number of Fixations	Total Time	Average
T	312 152 152 361 ST. PE(TE]RSBUR|G, No[v.) 2.—Th|e Admiralty...	4	1007	251
H	490 110 140 250 ST. PE[TERSBURG, [N)ov. 2.—|—The Ad[m)iralty..	4	1020	255
S	416 306 136 136 340 [ST.]PETER|S(B]URG, Nov|. 2.—(Th]e Admiralty..	5	1334	266
F	670 1116 646 ST. P|TERSBUR[G, Nov). 2.—The A[dm)iralty....	3	2432	810
M	811 148 140 171 351 163 273 (S[T. |PE]TER|SBUR[G,)No[v.) 2.—[Th)e Ad|miralty	7	2057	293

2

Subjects		Number of Fixations	Total Time	Average
T	123 152 218 has telegraphe|d to the offi|cers of the |Baltic........	3	493	164
H	330 220 160 240 ha[s telegra)phed to th|e o(fficers] of the [Bal)tic.....	4	950	237
S	314 144 306 238 238 263 136 [) |has telegrap[he) (d] to th[e of)ficers| of th|e Baltic..	7	1639	234
F	291 179 382 587 529 |has telegr|aphe|d to the [o)fficers of the B|altic......	5	1968	393
M	216 195 234 226 171 179 78 |ha|s te'egraphed [to) the of|fice|rs of the |Ba[lt)ic....	7	1301	185

Figure 58. The vertical lines and brackets indicate the locations of the eye fixations. The numbers give the length of the fixations in 1000th of a second. After Dearborn ('06).

Figure 58 shows the locations of successive pauses, their lengths in 1000th of a second, and the distances between them as determined by Dearborn. The brackets and vertical lines indicate the locations of these fixations. Sometimes there is a slight movement after the fixation. The direction and distance of this movement is indicated by the parentheses. These regressive or corrective movements occur when the eyes have moved too far to the right. They also occur when the eyes make the sweep back to the beginning of the next line but have not gone far enough. These corrective movements come more often in long lines than in short ones. They are probably caused by the fact that the peripheral perception of the beginning of the line is not accurate. The exact location from the beginning and from the end of the line is gotten as a habit after several lines have been read if the lines are of uniform length. For this reason the lines should be of the same length, especially for beginners. The opposite condition, however, is often found in reading texts in which each sentence is printed as a paragraph and in which pictures are inserted in the margins. The latter is still more objectionable when the pictures are colored, since the reflex action of colors influences the position of the fixations. The numbers above the vertical lines, Figure 58, indicate in 1000th of a second, the length of each pause. It will be noticed that these vary considerably from time to time and that on the average they are approximately one-fifth of a second in length.

Schmidt found on the average nearly twice as many refixation movements in oral as in silent reading. These movements grow progressively fewer and shorter from year to year as children progress through the grades. Gray found frequent refixation movements very characteristic of slow readers. They were also much more numerous with all readers in extremely careful reading.

Dearborn found that, with newspaper lines, five persons made an average of 4.76 fixations per line, that the number of fixations tends to decrease and with it the range of apprehension tends to increase as a passage is read over and over again, and that the eyes finally develop considerable precision and accuracy in the number and the position of the fixations. He also found evidence for the belief that the eyes form a short-lived motor habit which determines the number and position of the fixations for each line in reading. The formation of this habit develops during the reading of the first three or four lines of print and seems to come much sooner for some persons than for others. The fast readers seem to form the motor habit more quickly than the slow ones. The decrease in the time is due to a shortening of the pause and a widening of the scope of attention, especially in the latter parts of the line.

The fixations may come at any point in the line, between words, at the beginning, at the end or in the middle of a word. "The first and last fixations generally fall within the edges of the line, that is, a little distance from the beginning or the end words." The first fixation in a line is usually longer than the other fixations.

There are fewer pauses in the re-reading of a given material, even after one month, and the number decreases somewhat as the selection is read repeatedly. Prepositional phrases, connectives, and substantives make the greatest demand upon perception due to the fact that a slight change in them affects the whole meaning of a sentence. Dashes, punctuation marks, and capitals in the middle of a line change the points of fixation, and disturb some readers more than others. Reading a foreign language or reading aloud requires longer fixation than reading one's native tongue or reading silently. Dearborn obtained records of three boys aged 9, 10, and 11 years, in the third, fourth, and fifth grades respectively, and found that the oldest one approached very closely to the conditions of fixation of adults, while the younger ones made many more fixations or pauses. Fatigue resulting from the use of the eyes during an entire day decreases the rate of reading by about one-tenth.

Schmidt ('17) made an investigation of the eye movements, in oral and silent reading, of eighty individuals including elementary, high school, and university students. He found that they made 1.6 more pauses per line in oral than in silent reading, that the average duration per pause was from 20 to 27% longer, and that the perception time was from 44 to 64% longer.

By an ingenious combination of camera and phonograph, Gray determined accurately the relation of the eye to the voice in oral reading. He found that the eye always precedes the voice, with some subjects as much as four words and with others much less. As a general rule a wide eye-voice span was associated with fluent reading.

(5) Transmission of visual nerve impulses to the visual area of the brain. Little can be said about this process. All that we know about it is that the velocity of nerve impulses varies in different persons, and may be slow or rapid in slow or rapid readers, but we do not know definitely. Experimental work is necessary to determine to what extent it may be true.

(6) The arousal of association processes. A word gradually acquires meaning in the life of a child by its recurrence in numerous situations and by the connection of the particular significance or experience with the word. Thus, the child hears the word "chair" in connection with a certain object on which he sits. He hears it also after a while in connection with other objects which look different from the one he is accustomed to using. But new associative bonds are formed and he knows in general what is meant when he hears the word "chair." Later on he is shown a certain combination of visual characters and is told the word "chair." A new bond is then formed between the sound and the visual stimulus of the word "chair." An important part of the act of

reading consists, therefore, in the arousal of associative bonds as soon as the visual nerve impulses are brought in from the eye to the visual brain areas. The quickness of reading depends, no doubt, upon the rapidity with which these incoming stimuli arouse associated bonds and thereby give meaning to the word.

That the rate of reading depends upon the rapidity with which the visual stimuli are interpreted is shown by such experiments as Huey made in which he showed that it takes about twice as long to read nonsense as sense material. He also found that words in context give fewer associations than words out of context. When the words were shown in context the associations would sometimes go ahead of the amount seen and would anticipate what was coming.

Hamilton ('07) attempted to determine the part played by context in reading by ascertaining the time required to perceive isolated words and words in various relations, such as in paragraphs, in miscellaneous sentences, and in miscellaneous phrases. His results, expressed in terms of seconds per word required for the reading, are as follows:

TABLE 80. After Hamilton

	Paragraphs	Sentences	Phrases	Words
Mean	.429	.456	.466	.660
M. V.	.047	.047	.041	.078
Per cent using paragraph as base	100%	94%	92%	65%

Isolated words apparently require a distinctly longer time for perception than words in context.

Hamilton also found that the first part of a word is more important in giving the meaning of a word than the latter part is, that the marginal impression at the right aids in the perception of the words seen in the next eye fixation, and that the upper and lower marginal impressions probably do not aid in perception. He also thinks it probable that the interpretation takes place during the rest period, that is, during the movement of the eye.

The central neural and mental activity of putting meaning into the incoming impressions is probably the most important step in the whole reading process. The important part of reading really is the reading of meaning into words. Judd, McAllister, and Steele believe that the essential factors controlling the rate of reading are central rather than peripheral and that it is a matter of assimilation rather than a matter of getting material into the brain. Ruediger has a similar opinion:

"In reading, a similar reinstatement of experience takes place as in thought or in oral communication. The printed symbol arouses the meaning that has through education and experience become connected with it. It is to the rapidity with which this meaning is aroused that we have to look for the cause of the differences in reading rate.

"Reading rate may then be taken to depend chiefly upon the rapidity with which meaning is aroused in the mind after the symbol is seen. This, in turn, is in the main dependent upon the person's native brain inertia."

(7), (8), and (9) Transmission of nerve impulses from the visual center to the motor speech centers and thence to the speech organs. The motor speech centers concerned in the control of the speech organs are highly specialized. In right-handed persons, they are located in the left hemisphere of the brain in the region of the fissure of Rolando. These processes obviously will be involved in oral reading and in speech, but they are also active in silent reading in the form of incipient speech movements, particularly in the tongue and, to a less extent, in other speech parts. They are active in the same manner as in speaking, only on a much smaller scale. To what extent this inner speech is an important part in the reading process is somewhat uncertain. It is believed by some investigators to be an important agency for maintaining the continuity of the thought aroused by the successive visual glimpses of the printed line.

Disturbances in the motor speech centers are known under the names of various types of aphasia which will not be considered here. But even in normal persons the rapidity and facility with which the neural centers act may have an effect upon the efficiency of reading.

Much valuable information could undoubtedly be obtained from a careful laboratory study of individuals with language defects, such as persons who have difficulty in learning to read or difficulty in the speech functions proper. The following case may be of interest in this connection. A boy, 17 years of age, in the second year of high school, had always had considerable difficulty in reading. His reading was extremely slow. He was finally sent to a private teacher of public speaking with the hope that his reading ability might be improved. After several months of work with this teacher, apparently no progress was noticeable. He was then brought to the psychological laboratory for examination to see, if possible, wherein his difficulty lay. As careful tests as possible were made to discover in which of the various steps in the reading process the difficulty might be. These successive processes were tested with such means as were available or could be devised, beginning with the first, the reception of the stimuli upon the retina. Inquiry showed that his eyes had been carefully examined and found to be only very slightly defective. Glasses had been supplied as his father was a physician. Tests were next made to determine the range of distinct vision by observing how large an area he could see distinctly at one time. Next, his span of attention in apprehending visual stimuli was measured. This was done by rapid exposures with a tachistoscope to determine the number of words or pictures of objects he could apprehend at one time. In the next place, tests were made to determine the control of his speech organs. It was thought that possibly his difficulty in reading might be in an inability to control rapidly his speech organs. Tests were made by having him repeat statements from memory as rapidly as he could and also by having him repeat short sentences from dictation. All these tests indicated that these various functions were normal. His eyes had only slight visual defects, his field of distinct vision was apparently of normal size, his span of attention likewise was as large as that of a normal individual, and the control of his speech organs was rapid and accurate. The inference by a process of elimination was that the chief difficulty in his reading ability lay in the central assimilation or association processes. It seemed that visual impressions were brought into the brain centers with normal speed and facility, but that there was, for some reason, extreme slowness in the mental interpretation of these stimuli. A test in silent reading and in oral reading showed him to be equally slow in both cases. His rate of reading was approximately that of a child at the end of the first grade, namely, 1.5 words per second. His comprehension of what he read was good. His general intelligence, as shown by other tests, was normal for his age. His work in other school subjects was satisfactory. It seemed, therefore, probable, although not absolutely certain in the absence of further tests, that his difficulty lay in the central interpretation processes. This case, similar to others, shows in an interesting manner the great intricacy of the reading functions, and the difficulty in determining precisely what may be the trouble in a child who is backward in a given school subject but normal or even superior in all others.

In order to ascertain more fully the part played by the various factors enumerated at the beginning of this chapter, a series of tests was carried out by A. D. Mueller under the direction of the author,[1] with 36 high school pupils. Speed and comprehension of reading were carefully measured by three different selections according to the writer's method. Then each of the capacities mentioned in Table 81 was measured, and the coefficients of correlation computed. These show that the two elements upon which rate and comprehension of reading probably depend most are the visual apprehension span for related and unrelated words and quickness of association, correlation Nos. 2, 3, 6 and 12. Apparently the amount apprehended and the quickness of giving meaning to the visual symbols are the most important factors in reading. Correlation No. 6 corroborates Nos. 2 and 3 since the number of eye fixations per line of print is probably inversely proportional to the amount apprehended at each fixation.

TABLE 81

Correlations between reading ability and various elements entering into reading ability

	Speed	Comprehension	Speed plus Comprehension
1. Visual attention span—letters.....	.40	.32	.41
2. " " " —unrelated words....	.64	.73	.70
3. " " " —related words	.70	.59	.69
4. Auditory attention span..	.30	.37	.34
5. Rapidity of voluntary eye movement....................... ..	.31	.40	.38
6. No. of eye fixations in reading	.54	.49	.67
7. Association with auditory stimuli..	.38	.28	.35
8. " " visual " ..	39	.56	.42
9. Quickness of articulation—alphabet	.34	.42	.38
10. " " " —rhyme..	62	.76	.63
11. " " " —dictation	33	.43	.38
12. Continuous association.........	.63	.46	.62

[1] Reported in an unpublished thesis in the library of the University of Wisconsin, 1918.

These results are supported completely by a very similar experiment of Gray's, although he computed no correlations. As in the author's experiment there was a close relation between reading ability and the range of visual apprehension for meaningful material but, "The difference between the span of attention of the good and poor readers disappears in a very large measure when nonsense syllables, digits, groups of the same digit or the *aussage* test are given." [1]

The Measurement of Efficiency in Reading

(1) **Essential Elements to be Measured.** In order to be able to determine the effectiveness of different methods and procedures in learning and teaching a given school subject, it is necessary to be able to measure with some degree of precision and objectivity the general ability in that function. In order to measure such abilities it is necessary, furthermore, to determine what the essential elements in the process are which ought to be measured. These would seem to be in the case of reading as follows:

(a) In silent reading, speed and comprehension should be measured. The prime purpose of reading is the comprehension of the thoughts presented upon the printed page. The second element is necessarily the speed at which the thought processes may be obtained.

(b) In oral reading the chief additional element to be measured, besides speed and comprehension, is the correctness of the pronunciation.

(2) **Method of Measurement.** Several types of tests have been developed for measuring efficiency in the various aspects of reading. These methods will simply be mentioned here without a detailed, critical discussion of their relative values or of their various techniques of administration and evaluation. Tests for speed and comprehension of silent reading, have been devised by Gray, Kelly, Starch ('15), Courtis, Brown, Fordyce, and others. The writer's test consists of a series of eight passages, one for each grade. It is administered by having the pupils read for 30 seconds and by having them write out as full an account as possible of the

[1] Gray draws the following conclusion which is at least worth consideration:—"This seems to prove that the inherent differences in mental capacity which exist between members of this group are not the causes of differences in the span of attention. It appears rather that differences in training—that is, acquired abilities to deal with meanings—are the source of differences in perceptual span."

portion read. Speed is expressed in terms of the number of words read per second; comprehension is expressed in terms of the number of words written after those words have been discarded which represent incorrect statements of thought, additions of ideas not found in the original passage, or repetitions of ideas previously stated. The tests by Gray, Courtis, Brown, and Fordyce measure rate and comprehension of reading substantially in the same manner except that in some of them comprehension is determined either partly or entirely by answers to questions, and in the fact that most of them employ fewer than eight test passages. For example, Gray uses three, Fordyce two, and Courtis one passage for all grades.

Thorndike's ('14) tests are designed to measure comprehension in reading, either of isolated words or of paragraphs. They are constructed on the scale principle. Similar tests have been prepared by Haggerty ('17). The Kansas silent reading test (Kelly) is designed to measure comprehension primarily and speed secondarily.

Gray ('17) has devised a test for oral reading consisting of a series of twelve graded passages which are used to determine the rate of oral reading and the number of errors in the pronunciation.[1]

(3) **Uses and Results of Measurements.** (a) The first obvious use of exact measurements of reading ability is a determination of the actual abilities of the individual pupils, classes, or schools. The first logical step in the management of any activity is a diagnosis of the conditions as they exist. Then, upon the basis of this diagnosis, it is possible to prescribe more intelligently what should be done. In order to be able to make definite comparisons of the abilities of the pupils, it is necessary to know the standard norms, or averages, of achievement in the various grades. For the writer's test these are as follows:

TABLE 82

Grades	1	2	3	4	5	6	7	8
Speed of reading (words per second)....	1.5	1.8	2.1	2.4	2.8	3.2	3.6	4.0
Comprehension (words written)......	15	20	24	28	33	38	45	50

[1] Further discussion of these tests may be found in the original sources in which these tests were reported, or in the writer's *Educational Measurements*, or in Monroe's *Educational Tests and Measurements*.

These measurements are shown graphically in Figure 59 and indicate the grade-to-grade progress in speed and comprehension, showing that there is a continuous improvement in both aspects from year to year. By reference to these norms it is possible to express definitely a given pupil's capacities in reading. It enables one to say, for example, that pupil A in the 4th grade has a reading ability equal to the average ability of pupils in the 6th grade, pupil B in the 4th grade has a reading ability equal to the average

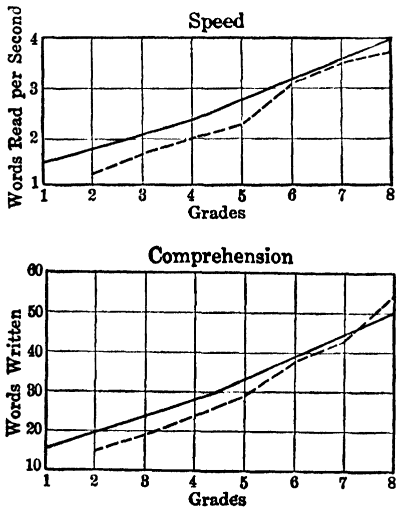

FIG. 59.—The continuous lines represent the standard attainments in reading. The broken lines represent the attainments in a certain school.

ability of pupils in the 4th grade, and pupil C has a reading ability equal to the average of pupils in the second grade, and so on. It is possible to describe precisely the ability of a pupil in relation to others because the terms in which his abilities are expressed are accurately defined. Measurements of this sort have disclosed enormous ranges of ability in the various school subjects as pointed out in a preceding chapter on individual differences. They have shown that at the present time pupils are not promoted according to ability, but rather according to the number of years they have attended school. Thus, for example, the pupils in a fifth grade range all the way in reading ability, from the second or third grade on up to the 8th grade. (See Figure 16 in Chapter III.)

Another interesting comparison at this juncture is the reading ability of the pupils of various ages in each grade. Such a study was made of the pupils in one school by the writer, and is shown in the following table:

TABLE 83

Showing the relation between age and attainment in reading

Grade Age	3 Speed Comp	4 Speed Comp	5 Speed Comp	6 Speed Comp	7 Speed Comp	8 Speed Comp.
7....	2.2 16.6					
8....	2.6 29.2	4.6 36.9				
9....	2.0 29.0	4.4 43.4	6.4 50.5			
10....		4.8 40.0	5.1 46.0	4.8 59.9		
11....		4.6 42.5	4.6 37.0	4.4 45.2	4.5 45.5	
12....		2.2 47.8	3.4 29.5	3.4 29.4	4.7 45.0	6.5 81.9
13...				2.3 28 0	3.4 39.8	5.4 64.1
14 .				2 8 16.0	4.6 52.5	4.5 52.0
15...				4 4 21.5		4.2 60.0

From the fourth grade on there is a fairly regular decrease in reading ability from the youngest to the oldest pupils in the same grade. The explanation for it is probably the fact that the stupid pupils are on the whole promoted too rapidly and the bright pupils too slowly.

(b) The establishment of accurate norms by means of definite measurements has meant, furthermore, a more precise estimate of definite aims of attainment. To say that a pupil at the end of the eighth grade should be able to read at the rate of four words per second and report a correct thought content expressed in at least 50 words, means something definite. The pupil, as well as the teacher, will know what each one means.

(c) The third and probably most important use of measurements of reading ability is their employment in the investigation of the factors and conditions affecting the learning and teaching of reading. In the long run the greatest good from tests of reading ability, or from tests of any school capacities, will come from their service in analyzing and measuring the potency of the numerous elements that enter into the acquisition of knowledge and skill in a school subject. Results on this phase of the subject will be summarized in the following section.

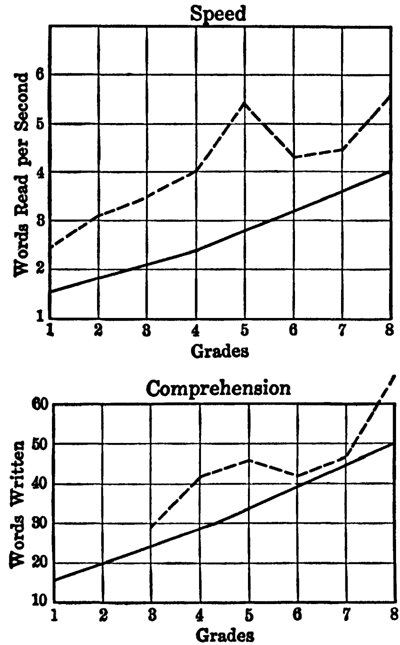

FIG. 60.—The continuous lines represent the standard attainments in reading. The broken lines represent the attainments in a certain school.

ECONOMIC PROCEDURE IN LEARNING TO READ

(1) **Differences among Schools and Classes.** The differences in achievement in reading among classes as a whole are enormous. Such differences are shown in Figures 59 and 60. Thus we see that the best classes in Figure 60 have scores approximately twice as high as the corresponding classes in Figure 59. These differences are too large to be explained by hereditary factors, but must be due chiefly to differences in learning and in method and spirit of teaching in these schools. What these differences in procedure of learning and teaching are must be discovered more fully in the future. But it is certain that some methods produce almost

double the achievement produced by others. Thus the school represented in Figure 60 is, grade for grade, as much as from 1 to 3 grades ahead of the average of schools generally, while the school presented in Figure 59 is from 1 to 2 grades below the average of schools generally. The former school has attained such proficiency in reading, probably chiefly because of the plan inaugurated by the principal which did away with nearly all word drills and phonics and placed all emphasis upon trying to read and upon reading as much as possible. He believes that a child learns to read by reading a great deal. He has described the method of teaching as follows:

"Since first grade reading is in the initial stage materially different from that of the succeeding grades, it is treated separately.

"The reading is begun in the first method reader by means of the word and sight method, and the phonics is carried on parallel with the reading, but in a separate period. No attempt is made in the early stages to correlate the phonics with the reading. During the entire work on the first book, the pupils are given words and sentence drills. Each lesson is read orally by every child after he has read it to himself. The time required to cover this book varies from two to four months, depending upon the ability of the child. When this book is finished, the child has a working vocabulary, and the method is changed to more extensive reading. The word drill is now omitted, and the lesson covering from five to fifteen pages is read over by the pupils, they or the teacher pronouncing the difficult words. Next the pupils read the lesson over silently, and the following day they are called upon to read a page each without assistance.

"When the seventh unit reader, at about the close of the sixth month, is in the hands of the pupils, the group reading is begun. The pupils are seated in groups of two, and each group is provided with an interesting reader. One pupil in each group reads aloud; when the first pupil has read three or four pages, the second pupil reads an equal number, then the first continues. In this way they alternate until the time is up. A fluent reader is often placed with a poor one for the purpose of assistance. At the close of a ten or fifteen minute period the place where the reading is discontinued is marked by a paper, and the books are laid aside until the next group reading period. Group reading is not, as a rule, practised oftener than twice a week.

"Sectional Reading:—This is a phase of group reading. Each group of two or more stands before the class. The first pupil reads three or four pages orally, the second continues the reading for three or four pages, and so on. In this type of reading the teacher may question the child as to what he has read, the child may reproduce it, or some listening member may tell what he has heard.

"Silent Reading:—During the study hour or seat work period, silent reading is conducted by means of single copies of books containing interesting material. Every child is given a book, and he reads as many pages silently as he can during the period. When the silent reading period is finished, a mark is placed where the reading ended. At the next silent reading period the pupil continues his reading, and so the work progresses until he has finished the book. The silent period may be continued as long as the teacher wishes. A record is kept of each child's reading by checking off a book as soon as it is finished. All through the year, a unit book is used during the regular recitation period for the drill that is necessary.

"Books are placed in the hands of the children on the first day of school, and they are allowed to keep books at their desks to read in school or at home as they desire.

"With this system, each child can go his own gait, reading as many books in a year as he can. The best readers will read from thirty-five to forty books. The average is about twenty books each." [Report of Reading in Dodgeville (Wisconsin) Public Schools, by Supt. H. W. Kircher].

The efficiency obtained by this process of teaching reading is shown definitely by means of various tests that have been given in this school as shown in Table 84. The results obtained in this manner from encouraging pupils to read so extensively are further indicated by the large number of books read as shown at the bottom of the table:

TABLE 84

Attainments in reading in a certain school

GRADES	1	2	3	4	5	6	7	8	DATE OF TEST
Starch Test:									
Speed in words per minute	96	126	156	240	318	252	264	330	
Comprehension in words written			28	42	46	41	46	65	Feb. 1917
Per cent above June standard in speed	7	17	25	66	90	39	20	37	
Per cent above June standard in comprehension			16	50	40	8	2	30	
Kansas Silent Reading Test:									
Score for each grade			12	19	16.5	14.1	19.6	23	Jan. 1917
Per cent above May			100	98	22	7	22	22	
Fordyce Reading Test:									
Speed-words per minute				209	272	250	276	250	April 1917
Quality				76	75	50	50	70	
Efficiency				65	69	22	20	37	
Courtis reading Test:									
Words per minute			150	190	220	220	220	280	May 1917
Comprehension			75	91	90	92	95	95	
Questions answered			34	41	36	40	45	67	

GRADES	1	2	3	4	5	6	7	8
Maximum number of books read by any pupil	38	96	90	150	101	120	105	100
Minimum number of books read by any pupil	20	45	41	29	42	20	17	18
Average number of books read per pupil	31	65	63	80	77	47	42	55
Average age of pupils	6	7.1	8.3	9.2	10.3	11.4	12.2	13.1

(Dodgeville Report, page 11.)

(2) The Possibility of Improvement in Reading Ability. Huey and other investigators report that they have been able, by special effort and practice, to double their speed of reading. Miss Harriet O'Shea (under the direction of Professor Henmon) conducted an experiment with a group of high school pupils in which she attempted to determine to what extent the rate of silent reading could be increased by specific practice. Her plan was carried out by giving each pupil a book, usually of high literary quality, and asking him to spend 15 minutes each day in reading the book until it was finished and to keep account of the number of lines read. Her results showed a rather remarkable increase in the rate of reading from the beginning to the end of the book. Some pupils gained very little and a few others gained very rapidly, as shown in the following table:

TABLE 85

Pupils	Average No. of Lines Read in First Two 15-Min. Periods	Average No. of Lines Read in Last Two 15-Min. Periods	No. of Lines of Gain or Loss	Percentage of Gain or Loss
1	696	1445	749	107%
2	793	1364	571	72
3	560	1038	478	85
4	553	964	411	74
5	461	648	187	40
6	430	758	328	76
7	420	715	295	70
8	416	655	239	57
9	403	569	166	40
10	367	732	365	99
11	365	528	163	45
12	364	467	103	28
13	355	390	35	10
14	322	667	345	107
15	303	362	59	19
16	275	480	205	74
17	263	685	422	160
18	260	415	155	60
19	253	371	118	46
20	245	240	−5	−2
21	241	484	243	100
22	241	377	136	56
23	235	343	108	46
24	235	224	−11	−4
25	224	317	93	41
26	204	125	79	38
27	142	139	3	2

Such experiments ought to be repeated again with additional attention to the question whether or not comprehension improves in a proportionate manner, and also whether this gain in the speed of reading would carry over to reading in general. At any rate, the results are significant in showing such large gains after a relatively short period of practice and effort to improve the speed of reading.

Peters ('17) undertook an experiment to determine the effect of speed drills conducted regularly in connection with the reading work during a school year. The pupils in grades three to six in the public schools of Royersford, Pennsylvania, were divided into "drill" and "no-drill" sections. The manner of conducting the different sections is described as follows:

"The groups which were not to have the speed drills, and which were to be used as a basis for comparison with those which did have, were dealt with after the usual fashion in teaching reading. The other groups, in addition to their oral reading, were given daily speed drills, without, however, giving a total of any more time to their reading than the other group received. So far as feasible both groups were taught reading at about the same time of day, or else at equally desirable periods. They used the same books and the same degree of enthusiasm was expected to be put into both. The drills were, of course, conducted by the teacher in charge of the class, and ran from November 7th to June 2nd. They were on relatively easy reading matter, and mostly interesting narrative. They occupied ordinarily from five to ten minutes of the reading period. The group as a whole was told explicitly where to begin and how far to read, and were then all set to silent reading at the same time with the exhortation to see who could get it read first. After all, or nearly all, had finished someone was asked to tell the substance of what he had read. If, in this reproduction, he omitted anything he was questioned on it as a guarantee against skimming."

Tests were given to both groups at four points during the school year, comparing the drill groups with the no-drill groups as a base. From the first test to the last, the results showed a gain in speed of 18.7% and the trifling loss of 1.1% in quality of comprehension.

Freeman ('16) reports a series of tests made by K. D. Waldo on the possibility of increasing speed. The lower grades particularly made a very large gain in speed which was accompanied by a parallel gain in amount reproduced, as indicated in Table 86.

TABLE 86

Improvement in reading from September to March [1]

	Rate in Words per Minute	Amount Reproduced	Percentage of Correct Answers to Questions
Grade 3			
September	76.4	71.1	44.6
March	149.1	135.	44.
Per cent gain	72.7	63.9	.6
Grade 4			
September	92.7	133 8	56 7
March	163.3	212.9	60.9
Per cent gain	70.2	79 1	6 6
Grade 5			
September	113.	52.2	16.3
March	129 2	70.5	25.6
Per cent gain	16.2	18.3	9.8
Grade 6			
September	128.	52.1	27.1
March	130.1	85.3	35.
Per cent gain	2.1	33.7	8.
Grade 7			
September	122.7	75.6	42.7
March	142.8	125.5	48.3
Per cent gain	21.8	49.9	5.6
Grade 8			
September	147.2	116.5	55.3
March	158.9	179.6	62.5
Per cent gain	11.7	63 1	7.1

(3) **The Relation of Speed and Comprehension.** This question has been one of perennial interest, and a common misconception has been held by a great many people regarding the mutual relationship of these two aspects of reading ability. Many people believe that a rapid reader comprehends relatively little of what he reads and that a slow reader makes up for his slowness by a more thorough comprehension of content. In order to obtain some specific facts on this question, the writer obtained the results from a careful test of reading in an elementary school in Port Townsend, Washington, and divided the pupils of each grade into six groups according to their speed of reading, putting the slowest sixth together and the next sixth together, and so on to the last sixth, consisting

[1] From an unpublished master's thesis by K. D. Waldo, on file in the library of the University of Chicago.

of the most rapid readers. The results are exhibited in the following table. The first column gives the average number of words read by each group, the second column the average number of words written representing a correct report of the thought, the third column gives the speed of reading in terms of the number of words read per second, the fourth column gives the percentage of words read in relation to the number of words written.

TABLE 87

Relation between speed and comprehension

Words Read in 30 Seconds	Words Written	Speed per Second	Per Cent of Words Written of Words Read
36	16	1.2	46%
51	22	1.7	43
69	24	2.3	35
90	33	3.0	37
105	33	3.5	31
147	54	4.9	37

These results indicate in a striking manner that the rapid reader comprehends relatively almost as much out of what he reads as the slow reader, and, absolutely, he grasps nearly as many more ideas in a given period of time as is proportional to the extra ground covered. Specifically, the table shows that the average speed of reading of the slowest group was 1.2 words per second, whereas the speed of reading of the fastest group was 4.9 words per second. The percentage of comprehension in relation to the amount read was 46% for the slowest group and 37% for the fastest group. In other words, the percentage of comprehension is almost as large for the fast group as for the slow group; or, comparing the first and second columns, we note that the fast group read almost exactly four times as fast as the slowest group and wrote three and one-third times as much as the slowest group. In other words, the ratio of the speed of the reading between the fast and slow group is one to four, while the ratio of comprehension is one to three and one-third. The inference is then that the rapid reader derives relatively almost as much out of what he reads as the slow reader. Absolutely he obtains several times as many ideas. Concretely, the comparison may be made in still another way: Of two persons belonging respectively to groups one and six, each reading for one hour, the fast reader would cover four times as much ground and derive three and one-third times as many ideas as the slow reader. The fast reader, therefore, has an astounding advantage over the slow reader. These results consequently give no corroboration for the common belief that an inverse relation exists between speed and comprehension in the fast and the slow reader.

Similar results have been presented by Judd. These results are shown in the following diagram, Figure 61:

"For the purpose of this study of the relation between rate and quality, all of the individual records of Cleveland pupils were divided into classes. First the speed records were arranged in order from the most rapid to the slowest. The most rapid of these records were designated by the simple term 'rapid.' In this class of 'rapid' records were included the most rapid 25% of all the records. In like fashion the slowest 25% of all the records were set aside and designatea as 'slow.' This left half the records, or the middle 50%, which were designated as of 'medium speed.' In like manner the 25% of all records which were qualitatively the best were designated 'good'; the 25% which were qualitatively the worst were designated 'poor,' and the term 'medium' was applied to the middle 50%.

"It becomes a very simple matter to assign all records in each grade to the appropriate class and determine the percentage of the grade which falls into this class. Diagram 59 gives the results, the percentages being in each case the nearest whole number to the calculated figure, and the size of the circle being proportionate to the size of the class indicated.

"These figures serve to emphasize the fact that good readers are usually not slow and poor readers are usually not fast. It is evidently not safe to attempt to lay down any absolute rule. There are good readers who are slow. In some cases such readers may be temperamentally slow. But even making allowance for such individual peculiarities, the figures show that good reading and slow reading are not incompatible. In like manner there are a certain number of children who read rapidly but retain little of what they read. With the figures in hand a teacher can profitably study her class and determine somewhat more completely than it is possible to do for the whole school system what are the special explanations of each individual type of ability.

"For the purpose of this survey the general fact that high rate and good quality are commonly related, and that low rate and poor quality are commonly related, is of great importance."

King ('17) tested the reading ability of 94 college students, half of whom were asked to read slowly and carefully and half to read rapidly and carefully. They read for ten minutes and, by observing a clock, the fast group were to read twice as rapidly as the slow group. Comprehension of the material was tested by answers to questions. The results showed that the accuracy of comprehension of the fast group was 44.5% and that of the slow group was 53.3%. In another test the subjects were divided into groups of naturally fast and naturally slow readers as determined by a preliminary test. The returns showed that the fastest 25% of the group had a comprehension of 50.2%, the slowest 25% had a comprehension of 48%, and the middle 50% had a comprehension

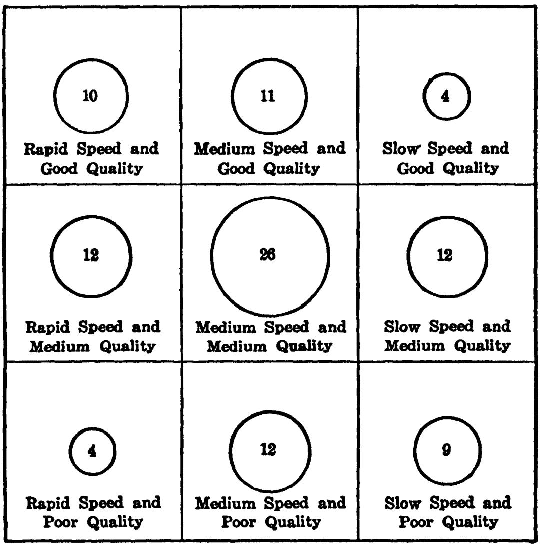

FIG. 61.—Relation between speed and quality of comprehension. After Judd' (16, p. 155).

of 46.5%. King interprets his results in favor of the slow readers. As a matter of fact they show, however, the same relation between speed and comprehension as found by other experiments. The differences in comprehension between the fast and the slow readers are very small, being slightly in favor of the latter group; but, when one remembers that the fast readers in the first experiment read twice as much text, the advantages are distinctly on the side of the fast readers. Whipple and Curtis ('17) in their study of skimming in reading also found that the slowest reader was the poorest reproducer and the best reproducer was one of the fastest readers.

(4) Relation between Oral and Silent Reading. Considerable attention has recently been given to the importance of relative stress upon oral as against silent reading or vice versa. The belief held by most of the investigators of this problem is that the schools have placed too much emphasis upon oral and not enough upon silent reading.

Superintendent Oberholtzer ('14) of Tulsa, Oklahoma, made a series of tests to ascertain the relative increase from grade to grade in the speed of silent and oral reading. Tests were given to 1,800 pupils. The following figures give the average speed of oral and silent reading in terms of words read per second:

TABLE 88

Speed in oral and silent reading. After Oberholtzer ('14)

Grade	Words Read per Second Oral	Silent
3	2.1	2.3
4	2.3	2.6
5	2.4	3.1
6	2.8	3.9
7	3.1	4.7
8	3.9	4.8

These results indicate that in the third grade the speed of oral and silent reading is very nearly identical, but that silent reading increases thereafter considerably faster from grade to grade, so that in the eighth grade the rate of silent reading is approximately one word per second, or about 25% faster than oral reading.

Mead ('15) tested 112 pupils in five classes in the sixth grade in both oral and silent reading. He made six tests, each two minutes long, and determined the speed by the number of lines read and the comprehension by the number of "points" reproduced in writing. His results are as follows:

TABLE 89

Relative ability in silent and oral reading. After Mead ('15)

	Av. No. Lines Read	Av. No Points Reproduced	Per Cent Reproduced of Amount Read
Silent reading	39.4	16.4	38.7
Oral reading	33.6	12.1	32.9

Each of the five classes did better in silent reading than in oral reading. Mead concludes: "From the results of these five classes

TABLE 90

Silent versus oral reading. After Mead ('17)

May, 1915 and 1916

(Six one-minute tests)

		Silent				Oral				
Grade	No.	Aver. Years	Aver. No Lines Rd.	Aver. No Pts. Repd.	Per Cent. Pts. Repd. of Pts. Read	Aver. No. Lines Rd.	Aver. No. Pts. Repd.	Per Cent. Pts. Repd. of Pts. Read	Per Cent. Difference Oral as Base: Total	Per Cent. Difference Oral as Base: Average
3	20	9.2	17.04	6.4	34.2	14.6	5.7	33.9	7.0	.35
4	20	10.1	27.0	7.2	23.9	17.4	8.7	35.0	−223.2	−11.1
5	18	11.1	15.7	7.9	45.57	13.1	5.86	39.38	111.4	6.19
5	21	11.9	25.5	11.7	43.1	20.3	8.5	36.2	143.8	6.8
5	21	11.4	25.2	9.9	35.7	18.38	6.6	30.19	114.	5.5
6	20	13.2	24.8	8.2	29.7	20.4	6.5	24.9	96.7	4.8
6	16	13.2	22.5	11.06	44.38	18.1	9.65	43.52	13.6	.85
6	20	11.7	29.5	14.2	45.1	22.5	11.3	46.0	−18.1	− .905
7	20	13.0	21.4	10.6	45.4	19.6	9.6	43.8	35.4	1.7
7	20	13.3	27.1	10.8	37.3	19.8	7.5	33.5	74.6	3.7
7	23	13.8	23.4	10.3	38.0	22.0	7.6	30.4	158.	6.9
8	20	15.0	31.1	22.5	66.0	24 2	15.8	51.4	293.0	14.6
8	20	13.4	26.0	16.7	56.9	23.0	14.5	54.3	52.4	2.6
8	20	14.1	22.7	13.7	58.0	22.8	11.5	45.3	254.0	12.7
8	20	13.8	25.2	15.6	55.1	20.2	11.7	50.5	93.1	4.6
8	21	14.0	25.6	13.3	46.7	23 0	10 1	39.8	146.0	6.95
10	20	16.0	21.2	13.8	56.6	20.3	11.8	49.7	139.9	6.99

The last column shows the per cent. of points each pupil did better by the silent method.

we are more convinced than ever that our schools devote altogether too much time to oral reading and too little to silent."

Within the following two years, Mead repeated the same tests with 340 pupils in grades three to ten, excepting the ninth, and obtained corroborative results. "Fifteen out of seventeen classes did better by the silent method of reading. Seventy per cent of the children taken separately did better by this method." The detailed facts are given in Table 90.

Pintner made eight tests, two minutes in length, with 23 pupils in the fourth grade and found the following results:

TABLE 91

Relative ability in silent and oral reading. After Pintner ('13)

	Av. No. Lines Read	Av. No Points Reproduced
Silent reading............	28.	18.
Oral reading..............	20.	15.

Thus it appears in all of the tests that silent reading has the lead over oral reading in both speed and comprehension. H. A. Brown ('16) and others believe that there should be no oral reading as such after the third grade, that silent reading should be emphasized instead, that above the third grade teaching to read should be teaching to study, and that there should be a great deal of spontaneous silent reading.

Superintendent Llewelyn ('16) of Mt. Vernon, Indiana, attempted to increase among his pupils the rate of reading and to stimulate interest in reading. He adopted the plan of giving oral reading three times per week instead of five times, of supplying a motive for silent reading by asking questions to test the knowledge of silent reading, of using the two extra recitation periods for live questions and discussions of what had been read, of stimulating the reading of library books and of having frequent book reviews. Each book was assigned to two pupils so as to make the discussion more interesting and lively.

No quantitative tests or comparisons were made, but the results reported were to the effect that the plan produced a love of reading, that teachers became more effective because they had to prepare for the giving of suggestions, that oral reading did not deteriorate and that reading was much more extensive as indicated by the fact that the class read about ten times as many books as before.

(5) **Phonics.** The tendency in recent years has been in the direction of less emphasis upon phonics and upon formal drills in general. Various investigators believe that extensive emphasis upon phonics and articulation in oral reading tends to establish slow habits of pronunciation and interferes with the proper development of speed in silent reading. Final experimental evidence on this question as on many other questions is lacking.

Currier and Duquid attempted to decide by experiment whether it was advantageous to teach phonics or not. No definite comparative tests, however, were made, but their impressions were that the phonic classes concentrated on the word and the sound at the expense of the sense, that their reading was less smooth and slower and that their ideas were confused. On the other hand, they reported that the no-phonics classes enjoyed their reading, that they read more swiftly, more expressively, and more for the sense of the material but that they did not read quite so accurately. The ability to attack new words was about the same. Experiments of this sort ought to be carried out more extensively and comparisons of the results should be made by means of more precise, quantitative measures.

(6) **Comparisons of General Methods of Teaching Reading.** Numerous methods of teaching reading have been advocated by publishers and educators, but no one knows with certainty the comparative merits of these methods nor which ones are most economically productive of the best development in reading ability.

Superintendent Harris ('16) of Dubuque, Iowa, in conjunction with H. W. Anderson, undertook an experiment to determine the relative merits of three systems of teaching reading. The teachers had felt for some time that they were not securing the results in reading that they might reasonably expect. The experiment is described thus:

"As a first move to remedy this unsatisfactory condition, the superintendent instituted a trial of the Beacon system at the school hereafter designated as School A, and of the Horace Mann system at school B. The present year was the second in which these systems have been thus used. The Aldine system was continued in use in the four other schools, C, D, E, F, mentioned in the report following. The teachers who worked with the Beacon system were enthusiastic in its favor; but their opinions, no matter how enthusiastically declared, were not sufficient to secure general agreement. With this state of affairs existing, the problem resolved itself into how to raise the question of the efficiency or worth of the various systems of primary reading out of the realm of mere opinion and place upon bed rock by scientific evaluation of results actually achieved.

"In order to accomplish this it was decided to test:

"1. The mechanics of oral reading in the second half of Grade I (1A) and in the two divisions of Grade II (2B and 2A).

"2. The silent reading in Grade II: for (a) rate; (b) comprehension.

"These tests were given in the following schools:

"A Where the Beacon system is used;
"B " " Horace Mann system is used;
"C D E F " " Aldine system is used.

"It was believed that these tests would show the results obtained by the Beacon system and the Horace Mann system during the first two years of their use, and offer an opportunity for comparison with each other and with the Aldine system."

Oral reading was tested with Gray's oral reading scale, and silent reading was tested by Starch's tests measuring the speed and comprehension of reading. The results obtained are given in the following tables.

TABLE 92

Results of the oral reading test

	IA	IIB	IIA	
A	36.5	52.9	55.7	Beacon
B	23.6	45.	50.2	Horace Mann
C	10.	20.3	42.8	Aldine
D	a: 0, b: 19.3	23.8	43.2	Aldine
E	8	40.3	44.	Aldine

"These scores seem to indicate that in each grade, School A, where the Beacon system is used, excels all other schools in the simple mechanics of oral reading. In fact, the scores of Grade IA at School A are better than the Grade IIB scores of both the C and D, and not particularly far behind the Grade IIA scores of the Aldine group of schools. The figures also indicate that the results at School B, where the Horace Mann system is used, are slightly better than at the schools where the Aldine system is employed.

"The results of the Oral Reading Tests seem to show conclusively that the pupils trained by the Beacon System are very greatly superior in the mechanics of oral reading to those trained under the Aldine or the Horace Mann systems of reading."

Results of the silent reading test

TABLE 93. Rate of reading

		A	A	B	C	D	E	F
IIB	Average	2.2	1 3	1.7	.8	.9	1.7	2.
	Median	2.	1.3	1.7	.8	.9	1.7	1.9
IIA	Average	3.	2.3	1.8	2.	1.7	1.8	2.3
	Median	2.9	2.2	2.	1.9	1.6	2.	2.2

"Table 93 shows the average and median rate of silent reading. It shows that in Grade IIB the Beacon pupils at School A read at the average rate of 2.2 words per second, while the non-Beacon group in the same class read at the rate of 1.3 words per second. The Aldine pupils read at the following rates: C, .8; D, .9; E, 1.7; F, 2. The Horace Mann pupils at School B read at the rate of 1.7 words per second. This shows that in this grade the Beacon pupils read .2 of a word faster than the nearest competitor (School F) and that they read more than twice as rapidly as two of the Aldine schools. The difference between the Beacon group of pupils and the best Aldine pupils is not significant, however.

"In Grade IIA the Beacon pupils at School A read at the rate of 3 words per second, while the non-Beacon pupils in the same class read at the rate of 2.3 words per second. The table shows that the rate of reading in Grade IIA is clearly faster at School A (Beacon pupils) than at any other school; the nearest competitors being School F and the non-Beacon group at School A, where the pupils read at the rate of 2.3 words per second."

TABLE 94. Comprehension of reading

		Non-Beacon	Beacon	H. M.			Aldine	
		A	A	B	C	D	E	F
IIB	Average	26.2	20.5	18.4	10.1	10.	20.6	17.
	Median	30.	18.1	20.	9.4	8.5	19.	17.5
IIA	Average	31.5	19.9	15.1	20.1	11.8	24.3	22.3
	Median	31.9	17.5	19.	19.9	8.	23.3	18.3

"Thus, the Beacon pupils in School A, Grade IIB, on the average reproduced 26.2 words, while the non-Beacon pupils in the same class reproduced 20.5 words. The nearest competitor to the Beacon group of pupils is the group at School E, which reproduced on the average 20.6 words—5.6 words behind the Beacon group. Schools C and D made the remarkably low grades of 10.1 and 10.

"In Grade IIA, the Beacon group of pupils reproduced 31.5 words, while the non-Beacon group in the same class reproduced only 19.9 words. The nearest competitor in the Aldine group of schools was School E, where the average number of words was 24.3. The results obtained through the Horace Mann system seem to be below the average, this class reproducing only 15.1 words.

"The results of the Silent Reading Tests seem to show conclusively that the pupils trained by the Beacon System are far superior to those trained under the Aldine or the Horace Mann systems of reading.

"It was realized that objections might be raised to the results herein shown, on the ground that the teacher rather than the system was the strong factor in the results. While it is highly improbable that out of six different corps of teachers, the group teaching the Beacon System would be uniformly better, in each section tested; yet, as an absolute check upon this phase of the matter, the Silent Reading Tests were given in Grade IIIB, which entered school before either the Beacon System or the Horace Mann System was placed on trial in any school and which therefore could not have had, in any one of the six schools, its initial training in either of the two systems named." (Starch's Reading Test, Series A, No. 3, was used.)

TABLE 95

Showing median rate and comprehension of silent reading in Grade IIIB

	A	B	C	D	E	F
Rate	2.9	1.8	1 9	1.9	2.	1.8
Comprehension	24.	27.5	12.3	29.	27.5	19.5

"This table shows that the pupils at School A read more rapidly than those at any other school, their median rate being 2.9 words per second, while the nearest approach to this rate was 2 words per second in School E. However, in comprehension three schools excel Grade IIIB at School A. Pupils at School A reproduced 24 words correctly, while those at Schools B, D, and E reproduced 27.5, 29, and 27.5 words respectively.

"Furthermore, in the tests made in Grade II, while the Beacon group at School A excelled all other groups in rate and comprehension of silent reading, several of the other schools excelled the non-Beacon group at School A in both these points. Thus, in the rate of silent reading, in Grade IIB, the non-Beacon group in School A was excelled by Schools B, E, and F, and in Grade IIA, the non-Beacon group at School A was equalled by School F; and in comprehension, the non-Beacon group at School A was excelled by School E in Grade IIB, and by Schools C, E, and F in Grade IIA.

"These facts show rather conclusively that it was not the superiority of the teaching which determined the results of the tests, since teachers in other schools than School A, both in Grade II and in Grade III, with pupils trained under the old systems, secured results as good as or better than those secured by the teachers at School A with pupils whose first training also had been under the old system."

Whether or not the results would be generally superior in other schools and under other teachers cannot be inferred perhaps from these results with complete finality. The experiment, however, is interesting and is here cited chiefly for the purpose of showing what should be done in the way of scientific comparisons and tests to determine the most proficient methods of learning and teaching reading.

Gray made a comparison of the attainments in reading in 44 schools, 26 of which had used the Aldine method, 17 the Ward method, and one a method of its own. The results showed no consistent or uniform superiority of one method over another. The average test scores were approximately the same.

Waldo ('15) compared the Howe system with the Ward system. The latter had been used up to the sixth grade in all the schools except one in which the former had been used. The tests were not carried out in a sufficiently careful manner to warrant reliable conclusions.

Hendricks made a comparison of schools in which the Rational method had been used with schools in which no special method had been used. He found the former schools superior. This does not necessarily prove the superiority of the Rational method but probably the advantage of well-organized systems over less well organized systems. The factors making for efficiency are so numerous and intricate that much more extensive and far more careful experiments will have to be made to demonstrate comparative values of the different systems or methods of teaching reading.

(7) **Suggestions for Improvement in Reading Ability.** Experimental results have brought about a radical shift in emphasis upon the aims to be accomplished in reading. In the first place, there has been a shift from emphasis upon oral reading to emphasis upon silent reading because facility in reading, in the sense of thought-getting, can be developed to a much higher degree of proficiency in silent reading and because nearly all the reading done by the average adult is silent reading. In the second place, there has been a distinct shift from emphasis upon slow reading to emphasis upon rapid reading because tests have shown that rapid reading does not mean a corresponding loss of thought, as assumed by many teachers, but, instead, rapid reading is accompanied on the whole by an almost equal ability in comprehension. The rapid reader will derive almost as many more ideas in a given period of time as is proportionate to the greater amount of text covered.

Pupils were formerly told that they must not read fast but that they should read slowly because they would then get the thought so much better. In the third place, there has come along with these two changes a shift from the mechanics of reading to the content of reading. Former aims of reading are fairly represented by the following answers, given by pupils who had finished the grammar school, in response to the question, What is your idea as to what the reading lessons were for? Some of the answers were: "To learn to pronounce," "To help us in reading before people," "Just to pass away the time," and "I thought it was to learn us to use better language." (Briggs '13).

Accordingly then the aim in reading to-day is the development of speed in reading and a parallel gain in thought-getting. Recognizing this change in aim, what definite suggestions can be made to facilitate improvement in these aspects of reading ability? Experimental results are as yet too few to make many specific recommendations with complete confidence. However, several important suggestions may be offered.

(a) As to the speed of reading: Force yourself to read more rapidly. Continuous effort and practice in this direction will very materially increase the rate of reading as shown by experiments. Probably most adults, as well as most children, read far more slowly than they are capable of reading. So far as we may judge on the basis of experimental investigations of the reading process, speed of reading depends chiefly upon the rapidity of the assimilation and upon the span of attention and less upon the other steps in the reading process enumerated at the beginning of this chapter. The rapid reader assimilates more swiftly and grasps more words at each fixation.

Forcing oneself to read more rapidly than one's customary rate will at first interfere with proper comprehension, but in the course of persistent practice the more rapid visual and mental activities will become habitual and the comprehension will probably then come up to its normal amount.

(b) As to comprehension of reading: Grasp the thought with concentrated attention. (1) Speeding up the rate of reading tends also to stimulate greater concentration of attention upon the whole thought content. That both speed and comprehension of reading may be very greatly improved by practice, by reading a great deal with the definite aim of improvement, is shown by such results as have been obtained in the Dodgeville schools and elsewhere. The remarkable rapidity of reading was accompanied by an equally remarkable ability in thought-getting.

(2) Stop frequently to recall the essential ideas read. Comprehension will be greatly assisted by stopping at short intervals and asking oneself the question, What have I really read? What are the essential ideas? This will not only stimulate reading for thought-getting but will also help to fix ideas in mind and to relate them to larger units.

(3) Acquire the habit of looking for the essential ideas. This is very important in efficient reading. Skill in so doing will greatly facilitate speed of reading by covering more ground and by knowing what may be read very hurriedly or even omitted.

(4) Tests at frequent intervals. Measurements by means of the standard reading tests or by means of improvised tests patterned after one or another of these testing plans should be given at frequent intervals for two reasons. In the first place, they will afford the pupil himself a definite basis for discovering his reading ability and, by keeping his own record from test to test, they will furnish a powerful stimulus to the pupil to surpass his own preceding attainments. This point was elaborated more fully in Chapter XI. In the second place, intensive tests with emphasis on both rate and comprehension will give the pupil practice in the phases of reading in which the school has in the past not furnished adequate training. The great emphasis upon oral reading has tended to instill slow habits of reading and placed the primary emphasis upon the mechanics of reading rather than upon thought-getting. Tests of reading ability, for this purpose, may be improvised and given as often as desirable by having the pupils in a class turn to a specified page, read with their maximum capacity for a limited interval, say half a minute, a minute, or several minutes, note the point of stopping, and then write a full account of the thought content, or answer questions. Such a procedure might profitably become a regular part of the instruction in reading.

CHAPTER XVII

HANDWRITING

Processes or Steps Involved in the Act of Writing

An analysis of the various steps involved in writing (or copying) similar to that made of the reading process reveals the following elements:

(1) Reception upon the retina of the form of the letters to be written.

(2) Transmission of the visual impressions from the retina to the visual centers of the brain.

(3) Recognition or perception of the letters through the visual and other association processes.

(4) Transmission of nerve impulses from the visual centers to the motor centers of the fingers, hand, and arm.

(5) Transmission of nerve impulses from the motor writing centers to the muscles of the fingers, hand and arm.

(6) Execution of the muscular movements involved in the writing act.

(7) Return kinæsthetic nerve impulses from these muscular movements back to the kinæsthetic centers in the brain and thence through steps (5) and (6) to help in correcting and controlling the writing movements.

(8) Return visual impressions of the letters or marks as actually executed back through steps (1), (2), (3), (4), (5), and (6), in helping to correct and control the writing movements.

What do we know concerning the manner of operation and the importance of each of these steps in the complete writing process? Steps (1), (2) and (3), while important when the child first learns to write, practically drop out in the skilled writer in whom the ideational and visual processes in the brain and the kinæsthetic sensations from the writing act itself serve directly to control steps (5) and (6). In the practiced writer they simply serve as a general control in securing the proper alinement, size, and spacing of the letters. The practiced writer can write about as well with his eyes closed as with them open. The chief difference is in such features as alinement, spacing and heaviness of strokes. The letters themselves can be formed practically as gracefully and as quickly with the eyes closed as with them open. Miss Downey ('08) found, in her experiment to determine the effect of different distractions upon the writing habit, that the visual factor has an obvious function in acquiring new coördinations but has little effect upon the fully formed habit. The visual perception of the form of letters, while important, is probably not as important in learning to write, even in the beginning, as it is in learning to read, because the writing act depends largely upon the development of muscular control. The visual perception of the form of the characters to be written and of those actually written must serve as a guide in the attempts at writing, but quickness in visual perception is not as important in the writing act as in the reading act because the writing act is much slower even in the skilled writer than the visual perception of the forms is, whereas reading depends directly upon the rapidity of visual perception. Then, also, the child has usually learned to recognize the forms of the letters when he begins to learn to write. At any rate its importance is rather secondary in the writing process.

The chief elements in the writing process are those connected with the steps (5) to (8), particularly those connected with the steps (5) and (6). Development of skill in any muscular movement which is not instinctive or at least not as mechanically precise as an instinct, proceeds by trial and error. Attempts are made at carrying out the desired movements. At first some of them succeed, most of them do not; and through continued trials the erroneous attempts decrease and the successful ones increase until perfect control is established. The adult scarcely realizes the utter lack of control in the early attempts on the part of the child in making the writing movements. The nearest approach that the adult can make toward realizing the actual difficulties of the child consists in such an experiment as the tracing of an outline as seen in a mirror. The spatial relations are so completely new and different that a person has little or no conception of the direction in which to move. Figure 62 shows a reproduction of the first tracing of a star outline as seen in a mirror. It does very little good to reason about it. One may think he is going to move in a certain direction but finds upon making the movement that he is going in an entirely different direction. The child learns to write very much after the same manner. He proceeds by trial and error. The successful movements somehow become more deeply fixed in the nervous connections and consequently more and more numerous. The correct movements finally become associated with the visual perception of form and direction so that the movement can be carried out at will with precision and grace.

It is a matter of common observation that a child beginning to write not only makes the movements very slowly but also with much excess pressure. The latter point was investigated by Meumann. He had children and adults write on a little platform which was so supported that the amounts and changes of pressure during the writing were transmitted to a light lever which traced them ac-

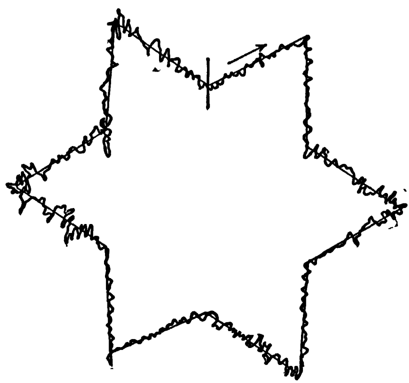

FIG. 62.—Record of tracing of a star-outline as seen in a mirror.

curately on a moving drum. It is natural to find that men tend to exert a greater pressure than women. Gross, working in Kraepelin's laboratory, found that men average nearly twice as great a maximum pressure as women. These records also serve to illustrate the fact that each letter has a fairly characteristic pressure rhythm. Freeman and others have shown by exact methods that the rate at which the writing point moves differs greatly in various parts of a given letter. In general the up and down strokes are the most rapid while the sharp turns and angles are the slowest. The speed curves of different persons writing the same letter consequently show striking similarities.

Sex Differences. The striking differences between the handwriting of different individuals has led many people to believe that there was a definite relation between the handwriting and the personality of the writer. The greater pressure exerted by men as a class suggests that the sex of the writer determines at least certain characteristics of the writing. This has been investigated independently by Binet, Downey, and Starch who find that untrained subjects can determine the sex of the writer in from 65 to 75% of the specimens (50% being pure chance). The author made a comparison of the writing ability of 2,113 boys and girls in the Madison schools and found the differences exhibited in Figures 63 and 64. It appears from these graphs that the median of the girls is above that of the boys in both speed and quality, but particularly in quality. The difference in speed is very slight.

Miss Downey ('10) attempted to determine sex differences in handwriting by selecting from envelopes by a chance method 100 samples of men's writing and 100 samples of women's writing and by asking thirteen persons to judge whether a given sample had been written by a man or by a woman. The thirteen persons made respectively the following percentages of correct judgments: 60, 60, 61, 64, 66, 66, 68, 68.5, 70, 70.5, 71.4, 71.5, 77.5. These show that on the average a judgment of sex as revealed in handwriting is correct 67% of the times, or two out of three times. Miss Downey also reports that the writing of the women showed less variability and more conventionality than that of the men. Those samples of writing by women which were called masculine were generally from persons accustomed to doing a great deal of writing.

Correlation of handwriting with other traits. A good deal of misleading character interpretation is based upon various features of handwriting. Experimental work needs to be done in this field, but it is probable that there is nothing to the claim of graphologists that handwriting reveals such traits as energy, clearness and simplicity, vanity, or self-consciousness.

Gesell ('06) reports that there is a close correlation between quality of writing and intellectual ability, but as a matter of fact his results show, as pointed out by Thorndike, a correlation of only about .30.

The author ('15) found for children a correlation of .31 between writing ability and general scholarship. Thorndike ('10) reports that for adults the correlation between writing and scholarship is zero. The probability is that the small correlation existing in the case of children is due to the fact of receiving instruction in writing and to the attention given to it. The better pupils do somewhat better in writing because they probably pay more attention or make more careful efforts. So far as adults are concerned, poor handwriting is no indication either of high or low intelligence, since the correlation is approximately zero.

Professional graphologists have claimed that a great number of specific traits of writing are determined by corresponding traits of character on the part of the writer. Thus the lines of writing of ambitious persons are supposed to slope upward from left to right. Hull ('19) investigated some of the more persistent claims by computing correlations between the traits of character of 17 university fraternity men, as judged by their fellows, and exact measures of samples of their writing. The correlations in each case were approximately zero, showing these claims to be entirely unfounded.

The Measurement of Efficiency in Writing

(a) *Essential elements to be measured.* The two important aspects of writing that must be measured are speed and quality, including under the latter legibility and form or beauty.

(b) *Methods of measurement.* Speed of writing is now generally measured in terms of the number of letters written per minute. Quality may be measured by either one of several scales, the Thorndike Scale ('10), the Ayres Scale ('12), the Starch Scale ('19), and others. The Thorndike Scale consists of a series of 18 steps or qualities of handwriting, each step consisting of one or more specimens of writing of the appropriate quality. Step zero represents an attempt at writing but as such is entirely illegible and devoid of beauty. Step 18 is a perfect copper plate specimen. The steps from 0 to 18 represent equal units of increase in quality. The Ayres Scale consists of 8 steps designated as 20, 30, up to 90. Each step contains three specimens of equal quality, a vertical, a medium, and a slant sample. The recent revision of the Ayres Scale, the Gettysburg edition, contains only a medium slant specimen for each step. The successive steps represent uniform increments of legibility in writing. The Starch Scale is composed of a series of 20 steps arranged in the order of merit or excellence of writing as judged by 400 persons. (See the original monographs for detailed description of the preparation of these scales.)

A sample of handwriting is measured by any one of the scales by putting it alongside the scale and determining which step it is most like in general quality. Speed and quality should ordinarily be measured simultaneously in the same sample because these two aspects of writing have a functional relationship. A test of writing should, therefore, be made by having the pupils write a short, simple sentence repeatedly as many times as they can in, say, two minutes, doing it as well as they can. Speed is then measured by the number of letters written per minute and quality is rated by one of the scales.

Freeman ('14) has prepared a set of five analytical scales for the purpose of rating handwriting from the standpoint of uniformity of slant, uniformity of alinement, quality of line, letter formation, and spacing. Each scale contains samples of three successive degrees

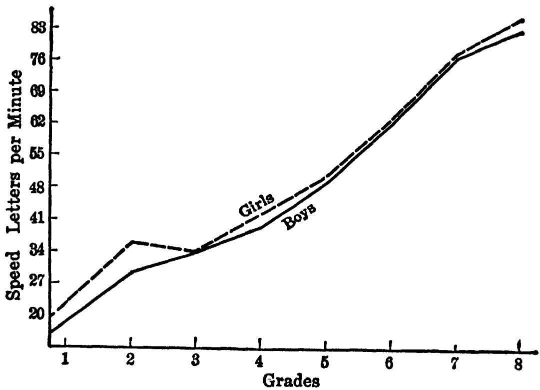

FIG. 63.—Sex difference in speed of writing. After Starch ('13).

of merit. These methods of rating ought to be useful in calling attention to defects in particular features of writing.

The A. N. Palmer Company has published a series of five or six samples of successive degrees of value for each grade. A percentage value for posture, movement, speed, and formation is given for each sample. The value of the sample as a whole is expressed by the average of these four estimates.

The Zaner and Bloser Company has also issued a set of specimens for evaluating handwriting consisting of a series for grades one and two, another series for grades three and four, and a third series for grammar and high school classes. Each series has a number of samples whose rating is expressed in terms of percentage values. Comments concerning the defects or excellencies are appended to the various specimens.

The manner in which the values of the samples were determined is not indicated for either the Palmer or the Zaner scales. This impairs their scientific value. From the practical standpoint they are commendable in that they suggest specific attention to, and evaluation of, important elements in handwriting.

(c) *Results and uses of measurements.* In general the results and uses of measurements in handwriting are the same as those pointed out for reading. It is possible by means of these measurements to determine more precisely the actual writing ability of a pupil, class, or school and to compare it with standard averages for cor-

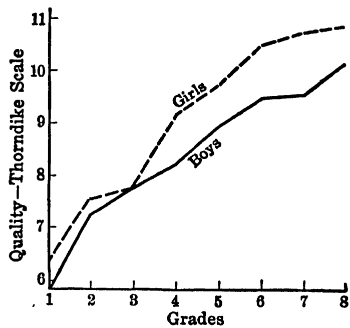

FIG. 64.—Sex differences in quality of writing. After Starch ('13, p. 461).

responding grades in schools generally. These standards of attainment for the ends of the respective school years are as follows:

TABLE 96. After Starch ('16, p. 83)

Standards of attainment in writing

GRADES	1	2	3	4	5	6	7	8
Speed (letters per minute)...	20	31	38	47	57	65	75	83
Quality (Thorndike scale)....	6.5	7.5	8.2	8.7	9.3	9.8	10.4	10.9
Quality (Ayres scale)........		27	33	37	43	47	53	57
Quality (Starch scale)........	8	9	9.7	10.3	10.9	11.4	12.0	12.5

By reference to these standards of attainment it is possible to define quite accurately the speed and quality of writing of a pupil or class by saying, for example, that a given pupil in the fifth grade is able to write 65 letters per minute at quality 9, Thorndike scale. The value of exact measurements of handwriting, as of any educational products, consists in the diagnosis of ability as it actually exists in different pupils and schools, in the measurement of the influence of different factors and conditions upon learning to write, and in the determination of the mutual relationships of various aspects of writing. A survey of our present knowledge concerning these matters will be given in a later section.

Economic Procedure in Learning to Write

What influence do the various factors, conditions, and methods in learning to write have in promoting or hindering the development of skill in handwriting? This question could be answered finally and fully only by the careful isolation of each factor under experimental conditions and by determining its effect upon the progress of learning to write. Substantial beginnings have been made in the direction of answering some of these questions, but little or nothing is known about most of them.

(1) Perception of the Forms Written or to Be Written. This topic requires consideration of two general questions: (a) What are the most advantageous conditions for the visual perception of the forms to be written? (b) What sort of form or model should be presented? The former question may be answered by the observation of certain obvious rules, namely, that the writing surface should be placed before the eyes at the proper distance, especially not too near, so as to avoid eye strain, and in a position directly in front of the eyes so that the points on the paper to be successively fixated may be at equal distances from both eyes, thus avoiding unequal accommodation in the two eyes. The paper should not be glazed, so that it will not produce a glare, and for young children the surface should be rather rough so that it will easily take pencil marks.

The second question is more complicated. The sort of models to be presented is obviously highly important since imitation, both voluntary and involuntary, probably plays a large part in the acquisition of writing skill. The author ('11) made an experiment in which he attempted to measure the unconscious effect of different models of writing upon the normal writing of adults. Four samples of writing were obtained from each of 106 university students. In order to avoid any suggestion of imitation, written rather than oral directions were given stating that they were to produce samples of their writing and that they should proceed at once to write the passages presented without further thought or questions. The four passages put before each person consisted of (1) a typewritten selection, (2) an extreme vertical model, (3) an extreme slanting model, and (4) a large model with many flourishes. The purpose of the typewritten passage was to obtain at the outset a sample of the normal writing of each person. The other three models were taken from school copy books.

After the experiment was finished, each person was asked whether he had tried purposely to imitate the various models. Three persons stated that they had intentionally modified their styles of writing. Their records were thrown out. The samples produced by the remaining 103 persons were carefully measured to ascertain their slant and size. Slant was measured by means of a specially prepared, transparent device with ruled lines for determining the angle of inclination of certain tall letters, such as l, f, and p, with the base line on which the words were written. Size was measured by determining the horizontal width of letters by measuring the length of words and dividing by the number of letters in the word.

These measurements showed that the average tendency for this group of persons was to make the letters distinctly more vertical when the vertical model was before them and more slanting when the slanting model was before them as compared with their normal styles of writing. They also tended to write slightly larger when the large model was before them. The amounts of these changes were as follows:

TABLE 97

Average inclination of l in the normal writing	65.1 degrees
Average inclination of l written from vertical copy	68.8 "
Average inclination of l written from slanting copy	61.5 "
Change from normal to vertical	3.7 "
Change from normal to slant	3.6 "
Total range of change	7.3 "
Average width of letters in normal writing	4.33 mm.
Average width of letters written from large model	4.85 mm.

When we realize that the handwriting of adults is a pretty firmly fixed habit, the amount of unconscious imitation is considerable, being a total of 7.3 degrees in slant and of .52 millemeters in width. We may infer that with children whose writing habit is in process of formation, the element of unconscious imitation plays a much larger part. Furthermore, it seems quite probable, although no experimental proof is at hand, that the style and quality of writing of the teacher distinctly influences the writing of the pupils, especially so because the writing done by the teacher in the presence of the pupils for the purpose of showing them how to write, is likely to be more efficacious in securing imitation than a static model in a copy-book would be. It would seem, therefore, highly imperative that every elementary school teacher should be a reasonably good writer.

In connection with the survey of penmanship in the Grand Rapids, Michigan, schools, Freeman (Judd, '16) reports that

"Grand Rapids adopted about five years ago a new system of penmanship. Up to that time the writing was not regarded as satisfactory. A part of the difficulty was thought to be due to the inability of the teachers themselves to write well enough to furnish a good example to the pupils. Accordingly, by action of the Board of Education, all teachers in the elementary schools were required, as a condition of promotion, to secure a Palmer certificate. This rule has been recently enforced with strictness and the writing in the schools is reported to be greatly improved."

Should the model presented to the pupils be vertical or slanting? Should it be plain or contain flourishes, decorative curves and shading? Should it be angular or rounding? The answers to these questions are at present largely matters of opinion and convenience rather than matters of scientific determination. Some years ago vertical writing came into general use because it was thought to be more legible and less productive of spinal curvature. But it has largely disappeared for the obvious reason that almost everyone naturally falls into the habit of writing a medium slant, no matter what style of writing was taught to him previously. The average slant for adults, as shown in Table 97 is about 65 degrees with the base line or about 25 degrees with the vertical line. Whether slant writing actually tends somewhat more to produce spinal curvature is doubtful. The difference in legibility between vertical writing and a medium slant writing is also probably very small. The letters should probably be of a medium slant and should be relatively plain and free from flourishes since these take time and add nothing to the general value of the writing, and finally, they should probably be moderately rounding because extreme roundedness is likely to reduce speed and extreme angularity is likely to reduce legibility.

Graves ('17) classified 604 samples of handwriting according to slant and then studied the speed and quality tendencies of the vertical, medium slant, and extreme slant group. The final averages of speed and quality are shown in the following table:

TABLE 98. After Graves

	Words Written in 5 Minutes	Quality (Ayres scale)
Vertical	91.6	57.98
Median slant	96.1	48.22
Extreme slant	101.7	43.58

There is revealed clearly a positive connection between slant and speed on the one hand and poor quality on the other. That is the "extreme slant" writers write more rapidly and more poorly than the "vertical" writers.

(2) **Length of Period of Practice.** What is the most productive practice period in learning to write? Even such a question as this, which is capable of definite experimental solution, has been answered only in part. The answer given by school programs in the time allotted therein for writing, is based largely on opinions instead of facts.

TABLE 99

Quality of handwriting at roughly the same rate in seven school systems. After Thorndike ('10)

Median results for eighth-grade pupils

System	A	B	C	D	E	F	G
At 20–29 words in 4 min.				14.5	13.0	15.4	14.8
At 30–39 words	11.5	11.3	11.6	12.3	12.3	14.5	14.2
At 40–49 words	11.5	12.0	12.0	11.8	12.3	14.4	15.3
At 50–59 words	11.5	11.6	11.1	11.1	11.6	13.0	11.7
At 60–69 words	10.3	11.8	11.5	11.3	11.6	13.6	
At 70–79 words	10.0	10.8	11.3				

Median results for seventh-grade pupils

System	A	B	C	D	E	F	G
At 10–19 words					13.3	14.5	13.5
At 20–29 words	12.3	13.3		13.0	13.6	14.2	13.0
At 30–39 words	11.0	11.8		12.3	13.3	14.2	13.0
At 40–49 words	11.0	11.8	11.3	11.7	11.0	13.3	11 8
At 50–59 words	10.3	11.4	11.1	11.0	11.8	13.0	
At 60–69 words	10.0	11.3	10.5	10.0	11.4	11.8	
At 70–79 words	9.8	9.8	9.9				

Systems A and B devote no time to writing as such in grades 7 and 8.
System C devotes 50–60 minutes weekly " " " " "
" D " 73–100 " " " " " " "
Systems E and G devote 60–90 " " " " " " "
System F devotes 75 " " " grade ".
" F " 30 " " " " 8.

Thorndike ('10) compared the writing in seven school systems as given in Table 99, and concluded that time was practically negligible. He says:

"What these facts do prove is: First, that at least three systems (C, D, and E) get little or no better results at a time cost of about 75 minutes a week than two systems (A and B) do at zero time-cost; second, that one system (F) at no greater time-cost than C, D, and E gets results about 25% better than they do; and third, that practice for quality may secure it only at the cost of speed. The teachers in A and B are better paid than those in the other cities, so that the success of these schools at no time-cost might not be generally attainable.

"Leaving F out of account, the differences of these school systems in the method of teaching handwriting, in the time devoted to it, and in the ideals of the system in respect to it are of inconsiderable influence upon efficiency. One makes its pupils write very well at very slow rates, the others vary a little in quality with small inverse variations in speed. On the whole, in spite of the achievement of system F, efficiency in handwriting seems, like spelling, and unlike arithmetic to be under present conditions not very much influenced by the management of the schools." (Thorndike, "Handwriting," p. 33.)

Freeman ('15) had writing tests made in 47 cities and then compared the attainment in these schools with the amount of time devoted to the writing-period in each school. His results are set forth in Figure 65. Each school is represented in the chart by a short vertical line. This line is placed at a position above the base line so that it represents the relative rank of that school in attainment in penmanship among the 47 schools.

These results are interesting and valuable, but it is questionable whether they prove that time makes no difference. The difficulty with a wholesale set of figures such as these is the impossibility of separating the various elements and determining their effects individually upon the ultimate attainment in writing. The schools which devoted 90 to 100 minutes per week to writing and obtained no better results than the schools which devoted 40 to 50 minutes per week, may contain other factors which kept their proficiency down, such as, poorer teaching, different classes of pupils, the quality of writing done in other subjects outside of the writing period, which probably has as much if not more influence upon proficiency in writing than the writing period itself, and so on. In fact, we might even imagine that if these same schools had devoted only 40 to 50 minutes per week, they might have been much worse in writing than they actually were. The real explanation may perhaps lie in the possibility that the schools having longer writing periods may not use the time to as good advantage as those having shorter periods. The latter, by virtue of having only a short time to devote to the subject, may work more intensely and profitably.

The surest way in which to measure the results obtained in different periods of practice in writing would be to split up a given group of pupils into several sections and to have each section devote a different amount of time to the writing period, say 10, 15, 20, and 25 minutes respectively. All should preferably be taught by the

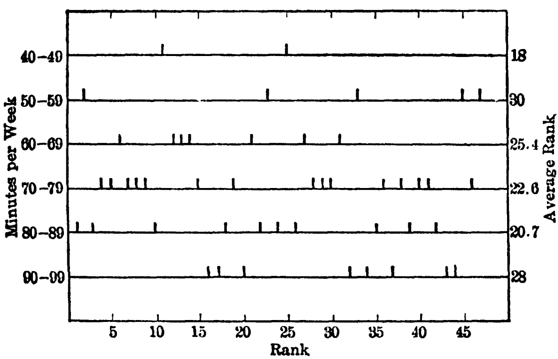

FIG. 65.—Relation between attainment in writing and time devoted to writing. After Freeman ('15).

same teacher. At any rate, all other elements should be kept as constant as possible. Then a comparison by special tests in speed and quality made at stated intervals would reveal the effect of time upon improvement and would show what period of time brought the optimum results.

The investigations by Thorndike and Freeman have been highly valuable in calling attention to this problem and in showing that some schools obtain as good results by devoting only half as much time as other schools obtain in double the amount of time. The general impression is that 15 minutes per day as a maximum is sufficiently long for the writing period and, under proper methods of instruction, can produce as high attainment in writing as the schools need to produce for all practical purposes. That there is an optimum length of the writing period beyond which the principle of diminishing returns operates is quite certain, as indicated by general learning experiments such as those cited in Chapter XI. What this optimum period for practice in writing is cannot at present be specified with certainty.

(3) **How Great Proficiency Should Be Attained.** Tests made in a large number of schools show that the average attainment at the end of the 8th grade is writing as good as quality 11, Thorndike Scale, or quality 60, Ayres Scale, at a speed of about 83 letters per minute. The same tests also show that many schools reach much higher proficiency than this and that in every school a considerable share of pupils far exceed the limits of 11 or 60 in quality and 83 in speed. Are these averages of attainment in quality and speed sufficiently high for practical purposes? And is it worth while to develop higher proficiency in writing than these averages represent?

In answer to the first question, Freeman ('15) made inquiries among business firms and found that the majority considered writing equal to quality 60, Ayres Scale, as sufficiently good for ordinary business purposes. It would seem then that the frequent criticism from business men who say that pupils coming to them from the public schools cannot write, is ill-founded, and based probably on the exceptions rather than on the majority of pupils, since about three-fourths of the pupils finishing the elementary schools can write better than quality 40 or 50, Ayres Scale, and one-half can write better than quality 60. The attainment of 11 or 60 in quality and of 83 letters per minute in speed reached by the average pupil upon completion of the elementary school is fully up to the average requirement of business. The criticism coming from business men is probably based upon the 20 or 25% of pupils finishing the 8th grade who fall below quality 40, and many times upon those who leave school before completing the 8th grade to seek business employment.

The second question is practically answered by the discussion of the first. It probably is not worth while to attempt to reach a proficiency in writing much higher than quality 11 to 12, Thorndike Scale, or 60 to 70, Ayres Scale. Such higher skill would be gotten by too great an expense of time and by too great a sacrifice of speed. Furthermore, the legibility of writing of qualities above these limits increases very little. The gain is chiefly in beauty. The time that would be required to reach these higher degrees of skill could be devoted to better advantage to other subjects, or to the learning of typewriting. Thorndike says:

"Considering the fact that above quality 11 there is very little difference in legibility, one is tempted to advocate the heresy that children are taught to write too well. I personally do advocate it. If school boards would furnish, for the use of children electing 'writing' as a study in the last two grammar grades, typewriting machines, I should certainly advise the transfer to typewriting of a child in these grades whose writing at 60 letters a minute consistently reaches quality 13. For, the amount of practice required to advance such a pupil to quality 16 at a rate of 75 letters a minute would much more than suffice to advance him to substantially errorless machine writing at that rate. The value now attached to the high qualities of handwriting is of course largely fictitious. Employers who can afford such high qualities of writing, buy machines to produce them. For writing cash checks, simple book entries, labels, and the like, a good plain hand or our quality 12 is entirely adequate. For attaining the higher qualities (15–18) the machine is a more economical tool than the pen, and in my opinion should be provided by those schools which require such qualities. Further, such qualities should, in my opinion, be required of children in the elementary schools, only when they have elected writing as a vocational subject. For the data from the adult women-teachers make it practically certain that ability to write above quality 14 will not be exercised in life except as a part of a clerical trade. If very, very few teachers find it worth while to maintain qualities above 14, it can hardly be supposed that it will be worth while for mechanics, house-keepers, farmers and dressmakers to do so." ('10, p. 37.)

(4) **Relation between Speed and Quality.** To what extent is speed of writing accompanied by good quality? Is there possibly an inverse relationship between the two? From general impressions we know that if we try to write unusually well we sacrifice speed and if we try to write unusually rapidly, we sacrifice quality. Is there any balance between these two elements?

Quite frequently teachers overemphasize either quality or speed, usually the former, at the expense of the latter. In Figure 66 the teachers in the 6th and 7th grades greatly over-emphasized quality so that the speed of writing was equal only to that of the average 3rd grade pupil. Definite tests and comparisons with standards will reveal to the teachers many such aberrations in emphasis.

Sackett had 36 university students write in their normal manner and immediately afterwards he had them write the same material with the knowledge that it was to be used as a writing test. He found that on the average the writing was about .5 letters per-second slower (original rate about 1.8 letters per second) while the quality gained 4 points on the Ayres scale.

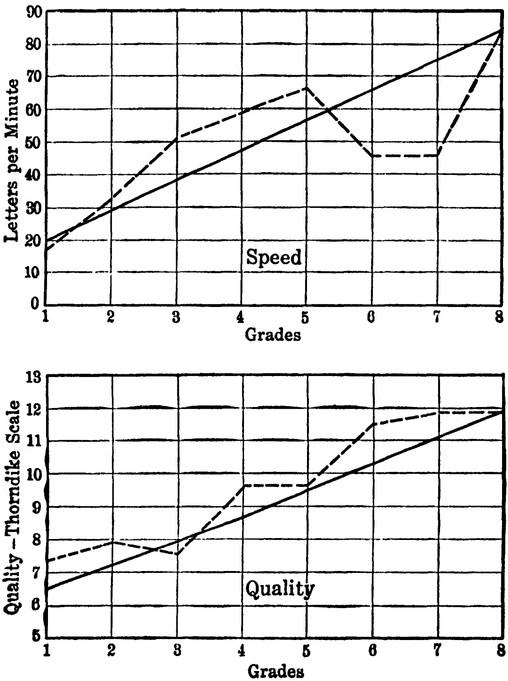

FIG. 66.—Average speed and quality of handwriting of the various grades in a given school. The broken lines represent the school. The continuous lines are the standard attainments.

Freeman ('14) made tests to determine what the efficiency would be when children were told to write (1) both rapidly and well, (2) as well as possible, and (3) as rapidly as possible. The results showed that trying to write well improved quality at the expense of speed. Quality improved 6.2% while speed dropped 3.7%. Trying to write rapidly increased speed by 27.2% but decreased quality 9.1%. Improvement in both speed and quality, however, can be obtained when instructions are given to stress both aspects. Apparently this is the preferable thing to do.

The author ('15) found in the case of 144 pupils the following correlations between various characteristics of handwriting:

Speed and quality	.10
Speed and legibility	.12
Quality and legibility	.34

Freeman also computed the correlation between speed and quality on the basis of the writing samples of pupils in Grades 4 to 8. These he found to be as follows:

Grade	IV	V	VI	VII	VIII
Correlation...	.08 (02)	−.10 (04)	−.14 (04)	−.37	−.15 (5)

These correlations are either zero or slightly positive or negative and mean that only to a very slight extent is the good writer extremely slow or the fast writer extremely poor.

Judd ('16), in the Cleveland *Survey*, has presented extensive data pertaining to this question:

"After determining the speed and quality of each specimen, it becomes possible to work out with great exactness the relation between these two characteristics. It is evident from ordinary experience that quality commonly deteriorates when speed is emphasized, and that speed is slow

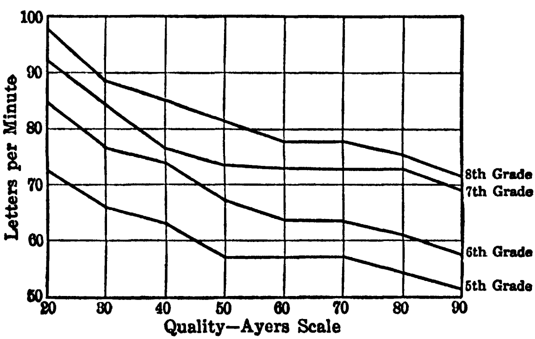

FIG. 67.—Average speed of handwriting at each quality of writing from 20 to 90, 10,528 cases from 5th, 6th, 7th and 8th grades. After Judd ('16, p. 72).

when one tries to write especially well. The school is constantly in the position of seeking some reasonable balance between speed and quality.

"Diagram 67 gives the facts for the 10,528 specimens carefully studied.

In the vertical axis of this diagram are represented the different speeds; in the horizontal axis are the various grades of quality. The results from each grade are represented separately. Thus, beginning at the extreme right end of the bottom line, we see from the diagram that for those writers in the fifth grade who show the highest quality (90) the rate is on the average 51 letters per minute. Advancing along the line toward the left, we find that those in the fifth grade who show a quality 80 have an average speed of 54 letters.

"The diagram shows that there is a general area between qualities 60 and 80, and between speeds 60 and 80, where all the grades above the fifth may be said to reach a level. Greater speed seems to be purchased at an undue sacrifice of quality, and higher quality seems to result in much slower speeds. We thus have in our results some indications as to the probable area within which teachers will find a desirable balance between speed and quality." (Pp. 70–71.)

(5) **Methods of Teaching Penmanship.** Experimental efforts have thus far not been directed very vigorously toward ascertaining the specific effect of different methods of learning to write upon rate of improvement in it. It seems, however, very certain that different procedures do produce enormously different results. This is amply shown by the wide differences in attainment of the various grades and schools, even in the same school system. Judd found, for example, in the Cleveland *Survey*, that the average of the best class was twice as proficient, either in speed or in quality, as the poorest class. The facts are shown in Figures 68 and 69:

"Diagram 68 shows the average results for the four upper grades in 36 schools. The figure is to be interpreted as follows: In the upper diagram, which gives the results for the fifth grades, there are numerous small squares, each representing a single fifth grade. In each square is a number showing the average number of letters written per minute in a grade. Thus in the square at the extreme left of the diagram is the number '39.' This means that the average number of letters written per minute by that fifth grade was 39. In the next vertical column of squares are numbers ranging from 42 to 49. These indicate that there were fifth grades showing each of the averages given.

"One of the most impressive facts which is brought out by this comparison is that the slowest fifth grade is only half as fast as its fastest fifth grade. Like statements can be made regarding the other grades. These wide differences cannot be attributed to any native characteristics which the children bring to the school. Such disparities might appear in individuals, but the figures report whole grades. All the fifth grades are going through the schools parallel with one another and are officially ranked as alike. The same statement can be made regarding the other

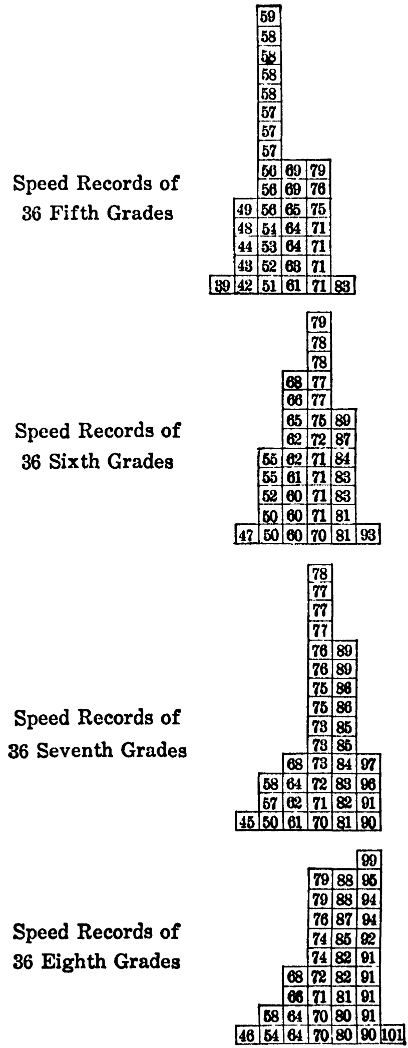

FIG. 68.—Distribution of grade averages in speed of writing. After Judd ('16, p. 64).

grades also. Perhaps the most obvious case is that of the eighth grade. Children will go out of the various eighth grades into high schools with the official assumption that they are equally well fitted for advanced

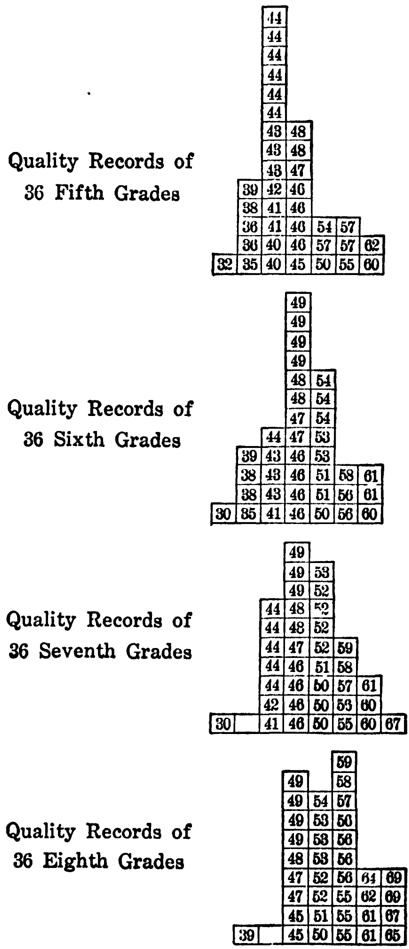

FIG. 69.—Distribution of grade averages in quality of writing as measured by the Ayres Scale. After Judd ('16, p. 64).

work, and yet one eighth grade averages only 46 letters a minute, and another averages 101. Is it not evident that there must be a difference in emphasis on speed in writing in different schools?" (Pp. 63–64.)

"Diagram 69 shows in a manner similar to that explained in the earlier paragraph on speed the results obtained from 36 schools. From the figure it will be seen that in quality, as in speed, the most striking variation exists between grades which are officially recognized as parallel. Furthermore, there is the same overlapping of grades, several of the fifth grades ranking higher than the average eighth grade." (P. 68.)

Here, as in so many other problems, specific experiments with conditions rigorously controlled should be carried out in order to determine the actual effect of each given element or method in teaching handwriting. Such experiments could be carried out by teaching parallel sections of a class according to different methods, after the general plan outlined under heading (2) of this section.

(6) **Factors Affecting the Execution of the Writing Movements.** The numerous conditions affecting favorably or unfavorably the execution of the many complex writing movements such as the position of the body, the position of the desk, the position of the arm, the position of the paper on the desk, the manner of holding the pen or pencil, and the like, are important problems concerning which likewise we have little scientific information. The procedures followed by teachers are based chiefly upon general observation and personal judgment. With regard to position of body, arm, paper, and desk, Freeman has suggested the relations shown in Figure 70. This relationship makes possible a natural straight front position before the desk with both arms on the desk, and with the paper tilted at an angle of about 30 degrees to the left. This position of the paper makes it possible for the hand to follow easily along the horizontal line of writing by simply turning the forearm on the point on which it rests on the edge of the desk as a pivot. Furthermore, the most natural direction of the up and down movement of the pen point is directly toward or away from the body, and with the paper in the position suggested, the writing will have a medium slant of about 25 to 30 degrees from the vertical.

(7) **Types of Writing Movements.** What sort of writing movements may be executed most economically in learning to write and ultimately in the perfected writing process? Considerable controversy has occurred over this question. Theoretically there

are at least three types of movements possible in the production of letters. One would be to do the writing entirely with finger and hand movements and to hold the arm absolutely quiet except for the turning of the arm from left to right to follow along the line of writing. The second would be to do the writing entirely with the arm movement and to hold the hand and fingers absolutely still. The third would be a combination of these two sets of movements in varying proportions. Advocates of various methods of teaching writing favor one or another type of movement. Probably

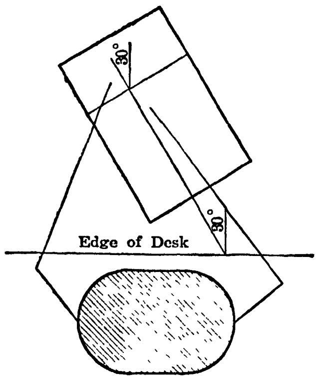

FIG. 70.—Position of pupil in relation to desk and paper. After Freeman ('14).

the best method is an appropriate combination as suggested in the third type of movement. Freeman states:

"The arm movement with rest—the so-called muscular movement—is an American discovery and has been vigorously exploited in commercial schools since the last quarter of the last century and more recently in certain systems of teaching in the public schools. It seems likely that within twenty-five years this form of writing will be practically universal in American schools. The chief advantages of the movement are two. In the first place, it is made with the fingers relatively relaxed, thus avoiding cramping. In the second place, the rolling movement of the arm upon the muscle pad of the forearm produces a firmness and evenness of line, and the fact that the movement is produced from a center at a considerable distance from the pen point results in regularity of slant.

"The contention that every detail of the letters shall be made by the movement of the arm while the fingers remain immobile is calculated to antagonize reasonable critics. The oscillation of the arm may well form the main basis for the upward and downward strokes of letters, but to require that every loop and turn and joining be produced by the movement of the arm as a whole, instead of the much more flexible hand and fingers, is to set up an artificial requirement and one which is not made in regard to other types of skilled movement.

"The form of movement, then, which best meets the requirements which may be laid down as the result of experiment and of practical experience is somewhat as follows: The hand and arm must be so adjusted that the hand progresses freely along the line during the formation of the letters and in the spaces between the words. The hand must rest upon some freely sliding point or points of contact such as the finger nails or the side of the little finger. When, on the contrary, the pen point is carried along from one letter to another by means of adjustments of the parts of the fingers and the hand, the hand continually gets into a cramped position.

"The movements of the arm and fingers should form a smooth and easy co-ordination in which there is a condition of flexibility in the whole member. The rotation of the arm upon the muscle pad of the forearm as a center carries the hand along, the upward and downward oscillatory movement forms the groundwork of the letter formation, and slight adjustments of the fingers complete the details of the letters. In addition to these chief elements of the movement, the wrist may rotate to the side to supplement the sideward movement of the arm, and the forearm may revolve upon its axis in the movement of pronation as a corrective to the increase in slant at the end of the line. There is no good reason for seeking to eliminate any of these component movements. Each has some part to play. Moreover, room must be left for individual differences in their relative prominence and manner of combination." ('14, pp. 93–96.)

(8) Movement Drills. Special drills in movements such as ovals, vertical movements progressing to the right, horizontal movements from left to right and from right to left, have been advocated by various systems of penmanship with the belief that a substantial amount of time given to such drills will establish good form and speed in writing. To what extent such formal drill or how much of it may actually be profitable, is open to question. It would be an experiment worth undertaking to teach three sections of a class of pupils for a year or more by giving to one section a considerable amount of such drill, to the second section none, but to devote instead the entire time to drill and practice in writing the letters themselves, and to the third section a combination of the two types of drill.

(9) **Correct Form in All Writing Done by the Pupils.** It is an elementary principle of habit formation that an act to be developed into a skillful habit, whenever carried out, should be performed correctly or at least as correctly as possible at that stage of learning. Otherwise the inaccurate and careless performance of the act tends to counteract the skill already achieved. It would seem, therefore, to be a highly desirable plan as an incentive to pupils to write at their best at all times, to base their marks in penmanship to the extent of one-half upon their work in the writing-period proper and to the extent of the other half upon the quality of writing in all other work submitted. One important reason why instruction in penmanship, spelling, oral and written composition does not carry over into the penmanship, spelling and composition generally is that pupils are not as careful in their ordinary writing, spelling and speaking by observing correct form as they are in the respective class periods devoted to these subjects. Telling the pupils that their final grades will be made up, half and half, as here suggested, will act as a remarkable incentive toward general improvement as shown by specific tests in the case of spelling, which will be discussed in the next chapter.

(10) **Analysis of Imperfections.** One of the important by-products of the experimental investigation of conditions and factors in the learning process is the fact that definite practice in a specific function consciously known to the learner greatly improves the function. Improvement in any type of skill takes place in many instances only when practice is squarely directed towards certain specific elements in the process. This is one reason why persons in laboratory experiments on learning make such enormous progress and why pupils in school make so little progress. The function to be trained in the one case is definitely and specifically known to the learner, whereas in the latter case, it is indefinite and largely unknown to the learner. It is not enough to say to a pupil, "You must write better," "write more like the copy," or "watch me; write as I do." The specific defects must be pointed out, recognized by the learner and then overcome by definite practice.

Freeman has pointed out five main types of defects or characteristics of handwriting: uniformity of slant, uniformity of alinement, quality of line, letter formation, and spacing. The scales that he has devised for rating handwriting from these five points of view may be used with advantage in discovering the specific defects in a given individual's writing and in centering definite attention and practice upon them. The score card for evaluating handwriting prepared by C. T. Gray calls attention to a similar set of elements. The methods of judging penmanship suggested by Palmer and by Zaner aim likewise to center attention upon defects and excellencies in various essential aspects of writing.

CHAPTER XVIII

SPELLING

Processes or Steps Involved in Spelling

The child learns to spell by seeing or hearing the letters of a given word, and by thinking, speaking, or writing them in the order in which they are seen or heard. Stated in more minute detail, the successive steps are substantially as follows:

(1) The reading of the word, that is, the sight, sound and pronunciation of the word as a whole which involves all the elements of the reading process and need not be enumerated here. (See Chapter XVI.) These are presupposed as the child usually has learned to read the word before he learns to spell it. At this point the successive steps in the spelling process as such begin.

(2) Reception upon the retina (or the ear) of the visual (or auditory) stimuli of the first letter of the word.

(3) Transmission of the visual (or auditory) impressions from the retina (or ear) to the visual (or auditory) centers of the brain.

(4) Arousal thereby of mental images and other associations of meaning.

(5) Transmission of the impulses from the visual (or auditory) centers to the motor-speech centers or to the motor-writing centers.

(6) Transmission of motor impulses from the latter to the speech-organs or to the writing-muscles. This occurs very probably even in the silent learning of spelling since silent reading is accompanied by the so-called inner speech.

(7) Execution of the speaking or writing movements in pronouncing or writing the letters.

(8) Return kinæsthetic impulses from the speech or writing muscles to the sensory centers and then to the motor speech or writing centers. This series of steps from (2) to (8) is then repeated for the second letter, for the third letter, and so on to the end of the word.

The steps here outlined are the ones involved in *learning* the spelling of a word. In the perfected process, however, steps (2) and (3) and possibly (4) drop out and step (5) is inaugurated directly either through step (1) or through the idea or image of the word to be spelled or written, and from then on the whole process of writing or spelling the word consists of a circular series of automatic connections between steps (6), (7), and (8) in which (8) for the first letter of a word acts as stimulus to step (6) for the second letter and so on for the succeeding letters of a given word. Step (8) of each letter always acts in turn as the stimulus for the series (6), (7), and (8) of the succeeding letter so that in the finished habit the mere pronunciation, sight, image or idea of the word automatically brings about the succeeding links involved in naming or writing the letters in correct order.

Economy in learning to spell consists largely in providing conditions under which the half dozen links here outlined may be established most easily, most quickly, and most permanently for the words whose spelling a child should know.

Little is known directly concerning the manner of operation of each of these factors. The most important step, if any one is more important than any other, possibly is number (8). This link determines what the next letter shall be in the automatic writing of a word. In the original learning of the spelling of a word, steps (2), (3), and (4), which together constitute the perception of the letters, are highly important since the establishment of the other links depends upon the accuracy with which the letters themselves are perceived or observed. It seems probable, although not certain in the absence of pertinent experimental data, that a considerable part of the difficulty of learning to spell, lies in the inaccurate observation of successive letters of a word. The awakening of mental images is probably very important, although our information as to the types of imagery concerned, the extent to which they are essential, and the methods of arousing them, is relatively unreliable.

The Measurement of Efficiency in Spelling

(1) **Methods of Measurement.** On the face of it, it would seem to be an easy matter to devise a definite and objective method of testing attainment in spelling. All that would seem to be necessary would be the selection of a series of words and the determination of the number or percentage of these words that a pupil or class can spell correctly. But a closer study of the possibility of measuring spelling ability reveals a number of complicated prob-

lems. What sort of words should be used as a spelling test? How should they be presented to the pupils? How should they be scored? Should any word be considered equal to any other word, or should different values be assigned to different words?

We shall not enter here into any critical discussion of the principles involved in the construction of spelling tests, nor into a consideration of the technique of administering and scoring them. Some of the methods of measuring spelling ability will be mentioned briefly.[1]

Up to 1913, tests of spelling ability were made either by series of arbitrarily selected words which were presented either as isolated words or as parts of dictated sentences, or by determining the percentage of misspelled words in spontaneously written compositions. Since 1913, several more or less scientific methods of measuring spelling ability have been devised.

The writer, ('15) prepared one test consisting of 6 lists of 100 words each by making a selection of words at certain intervals from the dictionary and then discarding all technical and obsolete words. The words in each list were then arranged in the order of length. Each list as a whole was found to be practically identical in difficulty with every other list. Average standards of attainment were then prepared for the various grades as shown in the following table, which gives the percentage of words of any one of the 6 lists spelled correctly at the ends of the respective years:

TABLE 99

Attainment in spelling. After Starch ('15)

Grades	1	2	3	4	5	6	7	8
Percentage of words correct	10	30	40	51	61	71	78	85

The author has more recently prepared a different method of testing spelling ability on the basis of the 2,626 most common words in the English language. This list is useful as a study list as well as a test list and will be distinctly more valuable because of this double purpose. The plan by which these words were selected and the method by which they are to be used will be described later in this chapter.

Ayres prepared a list of words consisting of 10 words for each

[1] For a detailed discussion, see the original monographs or the writer's *Educational Measurements*, or Monroe's *Educational Tests and Measurements*.

grade so selected on the basis of experiments that on the average 70% of the words for any given grade would be spelled correctly by the pupils of that grade. Later, Ayres ('15) prepared a very useful test consisting of the 1,000 most common words. These words are split up into 26 lists of varying length and so arranged that the words in any one list are of approximately equal difficulty and that the successive lists from 1 to 26 become harder and harder. The scale gives the average percentage of the words in the various lists that pupils can spell correctly in the different grades.

Buckingham ('13) prepared, on the basis of experiments, a list of 50 words carefully scaled in difficulty according to the percentage of pupils who could spell the words correctly.

(2) **Uses and Results.** All these scales have been found useful in measuring efficiency in spelling in different schools more accurately than was formerly possible, in determining individual differences in abilities among pupils, in ascertaining progress, and in comparing the effects of various factors in the learning and teaching of spelling. It is hoped that they will be still more useful in the future in discovering the most effective methods for acquiring proficiency in spelling.

The facts with regard to the enormous range of individual abilities in spelling and the consequent overlapping of the abilities of the pupils in the various grades are shown in Figure 18, Chapter III. The facts, as in other subjects, are astounding. The best pupil in the first grade spells as well as the poorest pupil in the eighth grade. Certainly some adjustment of the pupils should be made according to their capacities.

Spelling presents one of the more striking examples of mental sex differences found in educational psychology. Investigators uniformly report girls doing better than boys. Wallin found the girls averaging nearly 2% better than the boys in terms of his spelling lists. In a study previously quoted in Chapter XIV, Foster found that 238 girls and 256 boys, all university freshmen who were given a spelling test of 40 difficult words of Latin derivation, made respectively 76.6% and 68% successes. Sackett gave 24 words of Buckingham's spelling list to over 7,000 school children and found the girls about a half year of school progress ahead of the boys. Sears reports that a test composed of 70 words from the Ayres list given to nearly 13,000 children in Oakland, California, showed the girls superior to the boys from 2 to 6%, especially in the upper grades. He suggests that the average girl might be graduated from a half to a whole year earlier than the average boy.

The writer found the differences in one school as shown in Figure 71. The difference is in favor of the girls, particularly in the upper grades.

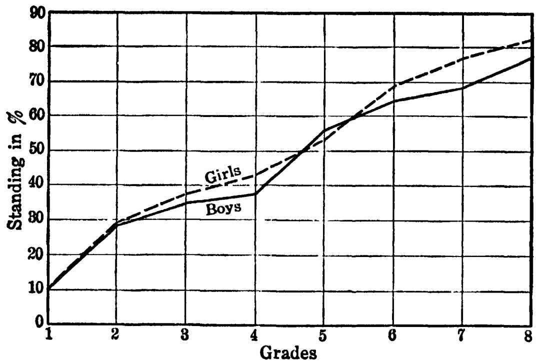

FIG. 71.—Sex differences in spelling as measured by the author's test.

ECONOMIC METHODS IN THE LEARNING AND TEACHING OF SPELLING

(1) **The Words to be Learned.** For a number of years the most important problem in the economy of learning to spell has been the question, What words should a child really learn to spell? The words in the spelling-books have for decades been selected mainly by the arm-chair method and have consisted to a large extent of rare and useless words. Swinton's *Speller*, published in 1872, states in the preface that "It omits the alphabet and the 'abab's' on the one hand, and on the other, quite a number of sesquipedalian words common to all old-time spelling books." It further urges as a vantage point, "The practical character of the work which aims to set forth, not the tens of thousands of long-tailed words in *osity* and *ation*, but the actual vocabulary used in speaking and writing." And yet the book contains in one lesson (page 143) such words as, lethean, pharisee, pentagon, pneumatics, theocracy, anathema, dysentery, etc. In another lesson (page 144) it contains oleaginous, farinaceous, argillaceous, lachrymose, sacerdotal, animadversion.

In view of this situation, recent years have brought forth a number of very extensive studies and tabulations to find the words which are most commonly used in writing by various classes of persons and to discover the frequency with which these words occur. The following are the chief tabulations thus made:

The Eldridge List. Mr. Eldridge ('11), a business man in Buffalo, New York, made a tabulation of 43,989 running words from four different newspapers in which he found 6,002 different words.

2,927	words	occurred	each,	once
1,079	"	"	"	twice
516	"	"	"	three times
294	"	"	"	four times
212	"	"	"	five times
151	"	"	"	six times
105	"	"	"	seven times
84	"	"	"	eight times
86	"	"	"	nine times
261	"	"	"	ten to nineteen times
238	"	"	"	twenty or more times

The Ayres List. Ayres ('13) of the Russell Sage Foundation tabulated 23,629 words from 2,000 letters, chiefly business letters, and found 2,001 different words.

The Jones List. Professor Jones ('13) of the University of South Dakota tabulated 15 million running words from 75,000 themes written by 1,050 pupils in grades two to eight and found 4,532 different words.

The Cook and O'Shea List. Cook ('14) tabulated 200,000 running words from the family correspondence of thirteen persons and found 5,200 different words.

These four tabulations represent four distinct fields of writing, each being the most extensive in its field; namely, journalistic, business, juvenile and private domestic vocabulary. One important type of vocabulary has never been tabulated, namely, the vocabulary of our best current literary writers. Children ought not to be confined to the words which they naturally use (Jones List), nor to adult business vocabulary (Ayres List), nor to newspaper vocabulary (Eldridge List), nor to the vocabulary of ordinary family correspondence (Cook List). An important point in learning to spell is to learn the meaning of words, especially of words whose use will enhance a person's vocabulary. Hence, the writer made a tabulation (Starch List) of the vocabulary of the best current literary authors. This tabulation is unpublished and on file in the Library of the University of Wisconsin.

The Starch List. The writer[1] tabulated some 40,000 running words, about 1,000 from each of forty authors in eleven current high grade magazines. This yielded 5,903 different words as follows:

3,111	words	occurred	each,	once
1,009	"	"	"	twice
512	"	"	"	three times
280	"	"	"	four times
189	"	"	"	five times
121	"	"	"	six times
97	"	"	"	seven times
82	"	"	"	eight times
53	"	"	"	nine times
225	"	"	"	ten to nineteen times
224	"	"	"	twenty or more times

From these five lists, words for spelling and testing purposes were selected according to the following plan: Every word occurring three or more times in the Starch List, every word occurring three or more times in the Eldridge List, every word occurring seven or more times in the Cook List, and every word in the Ayres 1,000 word list was selected if it also occurred in one other list including the Jones List. A word occurring three or more times in the Starch List or in the Eldridge List or seven or more times in the Cook List or any word occurring in the Ayres List was not included if it occurred only in the one list. To be included it had to occur at least once in one other list. This safeguarded against the inclusion of words confined to one type of vocabulary only. For example, the word "cupfuls" occurred twenty-one times in the Starch List but in no other list. Hence, it was excluded.

The reason for selecting words that occurred three or more times in the Eldridge List or in the Starch List was that the words found less frequently are so rare that they constitute a very small part of the running words of ordinary writing. This point may be shown most emphatically by the accompanying graph, Figure 72, on which the relative number of words of different frequencies is indicated. A remarkably close parallel exists between the Eldridge and the Starch Lists. The particular point to note in the graphs

[1] In coöperation with L. C. De Bruin. Reported in a thesis in the library of the University of Wisconsin, 1916.

is the fact that the sharp bend in both curves occurs between words whose frequency is between two and three. After three, the curve shoots up very rapidly. This same breaking point occurs in the Cook List between seven and eight. It is higher in this list because Cook tabulated a larger amount of writing. Words occurring three or more times in the Starch and Eldridge Lists constitute over nine-tenths of all running words.

This process of selection yielded 2,626 words. This number may seem small compared with the number of words in former spelling

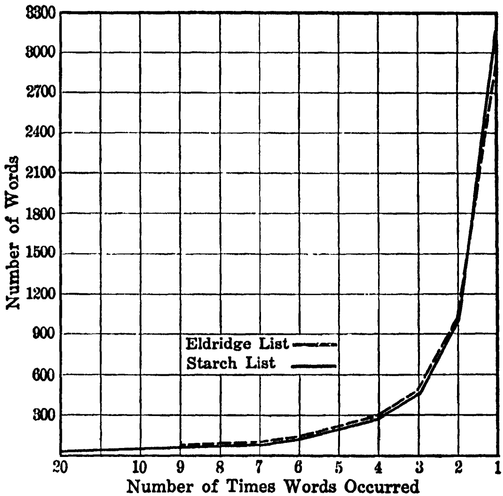

FIG. 72.—Number of words occurring with various frequencies in the Eldridge and Starch lists.

books and even in some contemporary spellers. But it is obvious that it is not only useless but wasteful of a pupil's time to learn words which he will never use in writing in all his years after school and at the same time neglect to master thoroughly the words which he will actually use. Spelling texts commonly contain from ten to fifteen thousand words. In fact, the number of words here presented, however, is even larger than that found in the special spelling lists prepared by many cities which often do not have more than from 1,500 to 2,000 words.

It seems, however, that a spelling list ought to include all words of reasonably common occurrence but not words of extremely uncommon occurrence. This number seems to be approximately 2,500 or 2,600. Another reason for limiting the spelling list to this number is the fact that the writing vocabulary of the average adult is not over 2,500 words and probably less for a great many people. Cook found that three out of his thirteen persons who had each furnished 40,000 running words (the largest number obtained from any one individual) of family correspondence had used a vocabulary of 2,575, 1,546, and 2,330 different words respectively. This would mean that a person could write 100 letters each 400 words in length, making a total of 40,000 running words, and not use more than 2,500 different words. If we may assume that the average man or woman, exclusive of persons whose occupation involves considerable writing, such as novelists, teachers, and journalists, writes one such letter a month, his entire correspondence for ten years would not involve more than 2,500 different words.

Incidentally we may also point out that this number furnishes enough words to supply two new words for each school day in grades two to eight, or 360 a year, counting 180 days to the school year, and 106 words for grade one. This is in accord with the practice of many school systems of teaching not more than two new words per day.

It is possible that such a list as that described above may need to be supplemented to meet the needs of certain sections of the population. Houser tabulated the words used by farmers in California in corresponding with the state department of agriculture. He found certain radical differences between this list thus obtained and that published by Ayres.

The Placement of the Words into the Various Grades. Can this be done on any scientific basis? Which of these 2,626 words should a pupil learn in each grade? There are three possible ways in which the words might be distributed into various grades:

(1) We might distribute the words according to their frequencies, putting the most frequent words in the lower grades and the less frequent words in the upper grades.

(2) We might put each word into the grade in which children first begin to use it rather frequently in their writing.

(3) We might put each word into the grade in which, according to the consensus of competent judges, such as teachers, it ought to be taught.

The placement of the words in the present list was made partly according to all three principles, but where discrepancies existed with reference to any given word final decision was made according to the third plan.

The 2,626 words were first arranged according to their frequencies of occurrence and all those words in the list which were also in the Jones List were then placed into the grades in which the Jones List shows them to be first used by children. A considerable number of the 2,626 words were not found in the Jones List. These words were placed into higher and higher grades according as they occurred less and less frequently.

After this task was completed the entire list of words was rechecked according to the third plan. Fifteen different lists of words which had been prepared by various cities or school systems for their own uses such as the Boston Minimum List, the Stockton List, the Santa Cruz List, the Chicago Speller, etc., were used. After each of the 2,626 words occurring in one or more of the lists, was written the number of the grade into which the word was placed by each list. An average of these placements was then obtained and accordingly the word was finally placed into its grade. For example, the word "flower" was placed by seven lists into different grades as follows: 4, 4, 2, 3, 2, 2, 4. This gives an average grade placement of 3.00. "Cough" was assigned by six lists to grades 5, 5, 3, 3, 2, and 4, with an average grade placement of 3.66. All of the 2,626 words were thus assigned with the exception of 126 words which did not occur in any of the lists employed and 178 words which were found in only one list. In order to make the grade placement of these 304 words with equal confidence, a group of seven experienced elementary teachers or supervisors were asked to assign each of these words to some one grade according to their best judgment. An average of these judgments was then obtained and the words were placed accordingly.

All of the 2,626 words were then assigned to the various grades according to the average grade placement as follows:

	All Words Whose Average Grade Placements were
In Grade One	From 1.00 to 2.00
" " Two	" 2.01 " 2.75
" " Three	" 2.76 " 3.28
" " Four	" 3.29 " 3.66
" " Five	" 3.67 " 4.71
" " Six	" 4.72 " 5.66
" " Seven	" 5.67 " 6.70
" " Eight	" 6.71 " 8.00

This process gave 360 words in each grade or two for each school day with the exception of grade one into which the remaining 106 easiest words fell.

Measuring the Attainment of the Pupils by Means of These Words. The problem that was next attacked was this: How many of the words of a given grade may we reasonably expect the pupils at the end of the year to be able to spell? In pursuit of an answer to this question, six lists of 60 words were selected from each grade list by taking, for a given list, every sixth word through the entire 360 words of a given grade. The words of the first grade were split up into two lists of 53 words each. These lists were then given as a special test at the end of the school year to approximately 7,000 pupils in 28 schools in 15 cities ranging in size from very small towns to a city as large as Seattle. The percentage of words of each grade spelled correctly by the pupils of that grade was as follows:

Grade	1	2	3	4	5	6	7	8
Av. percentage of words spelled correctly.......	56.2	63.6	77.0	80.4	83.9	85.0	82.9	80.7

Thus a very important advantage of these spelling lists[1] is that a school or teacher can at any time test the pupils and determine their efficiency by comparing them with the above standard averages. This can be done by selecting a list of 60 words from the 360 words for a given grade and giving them as a test. For example, if at any time the teacher of the fourth grade desires to compare the achievement of her pupils with the standard averages of other fourth grades in schools generally, all she needs to do is to turn to the words for the fourth grade and begin with any one of the first six words and then pick out every sixth word through the list. This will give a total of 60 words. At another time a different list of 60 words may be chosen in like manner. The same procedure may be followed in any other grade by using the words for that respective grade.

The important advantage of this plan is the fact that the same words which have been used as a study list may be used at any moment as a test list and comparisons may thus be made with the standard averages.

The averages here presented indicate that the words for all grades above the second are of approximately equal difficulty for the respective grades since they all are within a few points of 80%.

[1] Published in the author's Spelling Book, 1919, by Silver, Burdett & Co.

That the words in the second grade are not too difficult is shown by the fact that in several schools the percentages in this grade averaged as high as 90. The chief reason for the lower percentage in the first and second grades is the variation in amount of actual instruction in spelling given in different schools in these grades. The best school among the 22 averaged around 90% in every grade and the poorest school averaged around 60% in its various grades.

Aside from this list, numerous special lists have been prepared partly on the basis of vocabulary studies or partly on the basis of words commonly misspelled by pupils as reported by teachers. Illustrations of such lists are the Boston Minimum Spelling List (1915) consisting of 762 words, the Nicholson List, consisting of 3,070 words, prepared for the State of California, and the Chico (California) List, consisting of 3,470 words, prepared by Studley and Ware.

(2) **The Influence of Rules in Spelling.** Cook ('12) made a test with 50 words on 70 university freshmen and on 39 high school seniors and 30 high school freshmen. These 50 words were examples of seven rules with their exceptions. The university freshmen had had drill on spelling rules about seven months before the test and the high school classes had finished the study of rules six weeks before the test. The results are shown in the following table:

TABLE 100

Observance of the rules. After (Cook

RULE	Conscious of Rule While Writing				Unconscious of Rule While Writing				Combination of All Citing a Rule				Unable to Cite Any Rule			
	High School		University		High School		University		High School		University		High School		University	
	No. Students	Av. %	No. Students	Av. %	No. Students	Av. %	No. Students	Av. %	No. Students	Av. %	No. Students	Av. %	No. Students	Av. %	No. Students	Av. %
ie-ei.....	16	79	25	87	15	71	5	87	31	75	30	87	38	73	40	86
Final e ..	31	81	20	87	21	78	9	94	52	80	29	89	17	82	41	88
Final y.	11	74	18	94	18	67	13	95	29	70	31	94	40	73	39	91
Final con	15	78	32	88	27	72	2	87	42	74	34	88	27	75	36	84
Final ie	....	..	....	.	..	..	...	..	5	80	18	95	64	61	52	69

"In summary, it may be said that no one rule was quoted by as many as 50% of the university students, though more than half of them had memorized all these rules, and others besides, only the winter before; and many of the students had been over all the rules in the public school. A little less than half the high-school students had the courage to try to give the rules they had learned only six weeks previously. In the university group, those who gave some sort of rule to cover any part of the list of words, averaged 4% higher in general spelling efficiency than those who could not give any rule. So it is fair to assume that their better observance of the rules as shown by Table 100 is the result of their better spelling ability in general, and not to any conscious application of the rules as such. Not a single rule tested proved to be of real value, except the one for the last two words of the list—that relating to final ie." (After Cook.)

These conclusions are interesting and probably correct in their chief implications. One further point, however, ought to be considered. Inability to cite a rule or unawareness of its application does not necessarily prove the impotency of the rule. It might be possible that a rule played a part in the learning of words at the time the learning took place and then had been forgotten. A further investigation is necessary in which a comparison would be made between two groups, one of which had learned and applied rules while the other one had never had any contact with rules.

(3) **Length of Class Period.** Dr. J. M. Rice ('97) tested the spelling ability of about 33,000 pupils to ascertain the effects of different factors upon efficiency in the subject, such as methods of teaching, foreign parentage, home environment, amount of time devoted to spelling in the school program, and the like. His results with reference to the factor of time are presented in the table.

The results as they stand would seem to indicate that length of class period makes no difference in the ultimate achievement in spelling since schools devoting 10 or 15 minutes daily do as well as those devoting 50 minutes daily. Thus, City 15, School B, grade IVA, devoting only 15 minutes daily to spelling, made a record of 70.8 in the sentence test, whereas City 1, School B, devoting 50 minutes daily to spelling, made a record of only 61.8. Many other similar comparisons may be cited.

TABLE 101

Averages for individual schools in spelling. After Rice ('97)

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
		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	1	C
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	...	35	..	....	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	....	. ..	.	77.0		
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.

Averages for individual schools in spelling—Continued

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	98.7	..	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	98.2	..	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	....	.	73.2	.. .	..	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.

Cornman made a similar study and reports substantially the same situation. A sample of his results follows:

TABLE 102. After Cornman ('02)

Records of 1898

School	4th A Grade		5th Grade	
	Time Given to Spelling; Minutes per Week	Ability of the Class in Spelling	Time Given to Spelling; Minutes per Week	Ability of the Class in Spelling
A...........	50	67.0	50	68.0
C...........	75	75.5	60	72.5
E...........	100	76.4	80	83.5
F...........	100	65.0	100	57.4
G...........	100	66.0	100	76.2
D...........	120	65.0	100	66.0
I...........	150	70.3	100	67.1
H...........	200	76.3	160	82.0

The investigations of both Rice and Cornman are highly important but they do not afford conclusive evidence that time makes no difference. The fact that some schools devote two or three times as much time to spelling as other schools do and obtain thereby no better results, does not prove that time plays no part, for the reason that many other complicated factors enter, such as differences in teachers, method, spirit, and the like, to make the effect of any one factor unanalyzable. Indeed, we might infer that the lower schools, assuming all conditions the same, might have obtained even poorer results than they did if they had devoted to spelling only as much time as the better schools did.

In order to ascertain definitely the effect of different lengths of class periods upon spelling efficiency, it would be necessary to proceed in a more rigorously scientific manner rather than to make inferences on the basis of complicated, wholesale statistics. It would be distinctly worth while to undertake an experiment by teaching several equally able sections of a given class under as nearly identical conditions as possible, such as having the same teacher, text, method, and environment, and by varying only the time element so as to have, for example, a period of 15 minutes for one class, of 30 minutes for another, and of 45 minutes for still another. Comparisons by adequate tests at different times would yield conclusive evidence concerning the effect of time upon ultimate achievement. It would seem reasonable to anticipate that longer periods of equally good class instruction would produce greater efficiency. The fundamental problem, of course, would be the determination of the optimum length of the class period in relation to the desired efficiency in spelling.

(4) **Methods of Teaching Spelling.** Along with the question of time, Rice and Cornman were interested also in determining the influence of different methods of teaching. Rice does not present definite figures, but, on the basis of extensive inquiries among the teachers in the schools which he tested, he tried to ascertain the facts conerning methods of instruction and environmental conditions. His conclusion was as follows:

"In brief, 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. 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." (After Rice.)

Practically the same criticism, made in the preceding section concerning the factor of time, is pertinent here. The situation is too complex and the coöperating factors too numerous to make such inferences without a careful isolation of the individual elements and their separate effects. Each element should be subjected to an experimental procedure similar to the one outlined in the preceding sections.

Cornman went further than Rice by undertaking an experiment in which the spelling period was entirely eliminated from two schools in Philadelphia for a period of three years. He states:

"It was decided to abandon the use of the spelling book and home lessons in the subject, to omit also the period from the school programme which had been devoted to its study and recitation and to investigate the effect that the abstraction of these influences might produce upon the spelling of the pupils of the several school grades. Several methods of measuring results were devised which will be herein described and statistically reported upon." (After Cornman.)

On the basis of these tests, made at different intervals, Cornman found that the spelling ability of the two schools was almost as good as that of the other schools and that the pupils improved steadily in spelling, even though special instruction in them had been omitted. Some of his figures follow:

TABLE 103. After Cornman

Spelling ability measured by uniform examinations for all schools, given by the city superintendent.

	50 Schools Giving Specific Instruction in Spelling	2 Schools in Which for 3 Years no Specific Instruction in Spelling had been Given
7th grade	73.0	69.9
6th grade	70.4	65.1
5th grade	71.6	72.7
Average	71.7	69.2

The spelling ability of classes in the Northwest School who had for three years been without specific instruction in spelling compared with that of corresponding grades of previous years, who had had the full amount of drill in spelling.

	Tests of June, 1897	Tests of June, 1900
8th grade	99.4	99.8
7th grade	99.1	98.6
6th grade	96.75	99.0
5th grade	96.95	97.6

The spelling ability in a test in writing words in sentences of classes which had been without specific instruction in spelling for a year and of classes which had regular drill.

		Classes With Regular Drill Tests of June, 1897	Classes Without Regular Drill Tests of June, 1898	Difference
Northwest	8th grade	89.8	90.6	0.8
	7th "	86.1	78.7	−2.6
	6th A "	83.5	77.5	−6.0
	6th B "	72.7	71.7	−1.0
	5th A "	80.6	76.9	−3.7
	5th B "	78.8	79.2	0.4
	4th A "	75.1	80.7	5.6
	4th B "	85.8	85.3	−0.5
	3rd A "	86.5	70.4	−16.1
	3rd B "	57.8	57.7	−0.1
Agnew	4th A grade	76.8	82.0	5.2
	4th B "	82.5	83.7	1.2
	3rd A "	72.3	73.7	1.4
	3rd B "	66.1	67.7	1.6

That is, half of the classes without specific instruction did better and half of them did worse than the corresponding classes with specific instruction.

Cornman concludes:

"(6) The amount of time devoted to the specific spelling drill bears no discoverable relation to the result, the latter remaining practically constant after the elimination of the spelling drill from the school programme.

"(9) It is therefore advisable, in view of the economy of time, to rely upon the incidental teaching of spelling to produce a sufficiently high average result.

"(10) The average result is what can be and is attained, as shown by statistical evidence, by average pupils under teachers of average professional efficiency in classes of average size, i. e., in the elementary schools of this country as now organized. To remain strictly within the evidence gathered by this investigation, it must be admitted that there may be teachers of surpassing ability, who can obtain more than average results by the method of the specific spelling drill, and other teachers of meaner ability who need the drill to bring their pupils up to the level of this average result. It is claimed, however, that there is no evidence (whatever may be the wealth of opinion) to prove that such teachers exist or to show where they may be found. Moreover, the evidence which has been presented in this paper makes their existence at least improbable."

The extensive work of both Rice and Cornman has been very valuable in attacking a large educational problem by more exact methods and in showing that there is undoubtedly an enormous waste in the teaching of spelling, as there is probably in all school subjects; but the results are not final proof that length of time or methods of learning are negligible elements. Indeed both laboratory experiments in the learning process and the more recent and more carefully controlled tests in spelling itself indicate quite certainly that length of time and manner of learning do make important differences.

The author found that there are large differences in the average spelling attainment of the various classes of any given grade in the same school system. The results for the 10 schools of a Wisconsin city, Table 104, City J, were as follows:

TABLE 104

Differences in attainment in spelling in 18 Wisconsin cities as measured by the author's test

Grades		1	2	3	4	5	6	7	8
Standard Scores		10	30	40	51	61	71	78	85
City A					48 0	52.0			
C, School	1			36.5	53.5	61.7	68.8		
	2			34.6	47.1	60.7			
	3			38.0	43.7	57.1	68.8	76.2	75.2
F.		6.9	29.3	38.0	49.0	63.2	69.0	84.0	88.0
I, School	1			37.5	45.1	54.5	66 0		
	2			39.5	48.8	61.5	82.3	70.5	77.5
	3					52.8	64.1		
J, School	1		30.8	43.6	52.5	59.2	60.4	75 9	81.3
	2		35.0	34.0	47.0	62.0	79.0	80.0	83.0
	3	10.3	31.4	41.9	52.0	55.3	60.2	82.2	85.0
	4	6.1	19.0	32.0	47.7	58.0	68.0	75.0	77.0
	5	10.0	23.2	39.2	51.5	62.5	67.5	69.7	78.8
	6	8.1	25.1	36.1	45.8	53.2	62.7	74.0	79.7
	7	23.1	48.2	44.0	58.0	62.5	74.9	76.1	79.0
	8	7.4	25 8	37.7	44.7	55.9	67.3	75.0	77.9
	9	11.3	28.9	44.2	56.5	60.5	67.0	73.0	81.5
	10	10.2	35.2	43.2	54.0	55.7	64.8	71.3	82.2

Thus the lowest sixth grade made an average of 60.2 which is barely up to the standard of the fifth grade and the highest made an average of 79.0 which is above the standard of the eighth grade. These differences are actually quite large when we remember that they are grade averages and not scores of individual pupils. These differences must be due in a large measure to differences in methods of teaching.

Wallin attempted to make a comparison of the careful drill method, devised for the Cleveland schools by Assistant Superintendent W. E. Hicks, with the incidental method employed by Cornman in the two schools in Philadelphia from which the spelling period had been eliminated. Wallin reported results from the Cleveland schools distinctly superior to those in the Philadelphia schools.

TABLE 105. After Wallin ('11)

Combined averages for the composition and column tests, all schools

Grades	4	5	6	7	8	All Grades
Per cent	98.40	96.31	96.95	97.03	96.28	97.00

He states:

"First, the general spelling efficiency for all schools shown (97%) is striking. It is 12.6% higher than Kratz's results (84.4%, for the fourth to eighth grades, inclusive), 25% higher than Chancellor's (72%, for 10,000 pupils from the fourth to eighth grade), 25.48% higher than the results in Rice's column test (71.52%), which consisted of a list of dictated words, and 22.42% higher than the results from his sentence test, which consisted of dictated sentences containing 50 test words for the fourth and fifth grades, and 75 words for the sixth, seventh and eighth grades. It eclipses by 25.7% Cornman's average in three term examinations during three years for eighty Philadelphia schools (71.3%), and is 27% higher than the results of these examinations in his two experimental schools (70%), in which the spelling instruction was incidental. In four column tests given to these two schools from September to June and consisting of lists of fifty words differing from grade to grade, the averages were 33%, 49%, 50% and 50% respectively for one school, and 49%, 57%, 60% and 68% for the other; while the repetition of Rice's column and sentence tests gave an average efficiency in 1897 of 78.9% in one school, and 67.1% in the other, and in 1898, 73.1% and 65.4% respectively, for the *column test*. The corresponding averages for the *sentence* test were: 82.3% and 74.6% in 1897, and 76.5% and 77.9% in 1898. It will be observed that there was a loss of efficiency in 1898 except in the case of the sentence test for one school." (After Wallin.)

Miss Fulton reports an experiment in which an attempt was made to ascertain the effect of a specific drill in learning to spell. One hundred words, ten each day for two weeks, were taught by the following plan: 1. "Write word upon the board." 2. "Explain meaning." 3. "Children use the word in a sentence." 4. "Write word ten times while saying letters aloud at same time." 5. "Emphasize by intonation of voice or by colored chalk on blackboard the difficult part of words." At the end of the two weeks a test with the 100 words was given.

During the next two weeks another list of 100 words of similar difficulty was used with no directions except to "study the lesson." The children were "heard" but no special effort was made to teach the words. A test with these words also was then given. The results of the two plans as indicated by the tests given immediately and three weeks later were as follows:

TABLE 106. After Fulton ('14)

	With Drill	Three Weeks Later	Without Drill	Three Weeks Later
Average %	98	96	73	68
Daily results	98	..	78	..

(5) Laws of Association. Skill in spelling is primarily a matter of forming associative connections between certain arbitrary symbols arranged for the most part in arbitrary order. Economy in the learning of spelling reduces itself to this question: Under what conditions can these associations be made most quickly, most effectively, and most permanently? Of the four laws of association, frequency, vividness, primacy, and recency, the first two are most directly applicable. Obviously, frequent repetition is necessary to establish the connections. Frequent reviews, monthly, weekly, and possibly daily, are indispensable.

The law of vividness operates in numerous ways. This law states that other things being equal, the most vivid or intense association is most apt to be recalled. It may be made of service in two general ways: (1) By any device that will stimulate the clearest possible attention and interest on the part of the learner upon the spelling of words, or (2) by any device that will make conspicuous the particular part of any word that is most likely to be misspelled. These points will be considered more fully in the next three paragraphs.

(6) Careful Attention upon the Successive Letters of a Word. Pryor ('15) reports "an experiment to determine the value of 'spelling the word through' as an aid to learning. Two divisions of the fifth grade studied the same list. Conditions as to time, length of period, and the like were the same for both divisions. For one, emphasis was placed on observing carefully the order of letters while studying. Preliminary and final tests given to both divisions showed an advance from 50.55% to 83.39%, or an average gain of 32.84% for the division working under the usual conditions. The other division advanced from 48.58% to 89.14%, an average gain of 40.56%."

(7) Personal Incentive to Interest and Effort on the Part of the Learner. Aside from competitive contests in their various forms, there are at least two ways in which personal effort may be stimulated: (a) By having at regular intervals definite tests, preferably by means of some one of the standard spelling scales or tests so that each pupil may know his attainment from time to time and measure his progress. This plan will usually arouse very keen interest in the individual to surpass his own previous record. See Chapter XI. This is a form of incentive whose actual effectiveness has never been fully appreciated. (b) A second form of incentive is to arrange with the pupils that the grade in spelling will be determined half by the work in the spelling-class and half by their spelling efficiency in all written work. This plan was employed in a school in Potsdam, New York, with the result that the pupils developed a remarkable efficiency in spelling as shown in the accompanying graph, Figure 73. Each grade is approximately one entire grade ahead of the general average. This is one of the highest records found in any school measured by the writer's spelling test.

(8) Calling Special Attention to Difficult Parts of Words. One of the much needed investigations of spelling is a careful collection and classification of errors in the words used as spelling material.

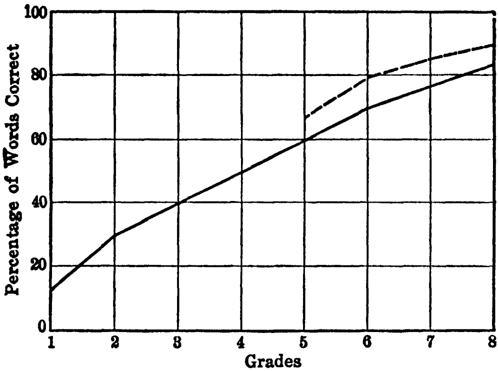

Fig. 73.—Averages in spelling in grades 5 to 8 in a certain school as measured by the author's test. The continuous line represents standard attainments.

Errors in spelling usually consist in certain letters or parts of words only. For example, the error in *receive* is usually in the two letters *e-i*, or in *separate*, in the *a* after the *p*. (Such a study is under way for the 2,626 most common words previously referred to.) By means of such a study the teacher would be able to anticipate the probable misspellings and to call especial attention to the letters likely to be missed. Such particular emphasis may be secured by asking the pupils to focalize the correct order of the letters which are usually confused, by writing those letters extra large, or by drawing a line around them, or by having them printed in larger or heavier type, and so on. It is quite imperative to anticipate the troublesome parts of words so as to forestall the formation of wrong connections. Instruction in spelling should consist in the teaching of correct forms rather than in the unteaching of incorrect forms.

(9) **Writing the Words.** We need to know how to spell words only in writing. It is a general law of association that connected bonds should be established in the manner and order in which they are to be used. It would seem, therefore, on *a priori* grounds that the associative bonds between the successive letters should be formed by exercise in the writing of the words so that the spelling may become automatic during the act of writing. Current emphasis upon writing the words in sentences is in the right direction. It would be an interesting experiment to teach two sections of a spelling-class by having in one section a great deal of writing of the words of the spelling lesson and by having little or no writing in the other section. The spelling ability of the two sections would have to be compared at various times by appropriate tests.

(10) **Context versus Column Spelling.** On the basis of the discussion in the preceding paragraph, it would seem obviously advantageous to have a great deal of writing of the words to be learned, especially in sentences. The belief is held by many that the ability to spell words in isolation does not insure ability to spell them in context. As a matter of fact, however, there is very little difference in spelling efficiency between these two situations. Wallin ('11) used the same test words both in a column test and in a dictated composition test and found an average spelling efficiency of 97.72%, in the column test, and 96.28% in the composition test, giving an advantage of only 1.44% to the former. Cook ('14) used 60 words in a column and in a composition test and likewise found only a very slight advantage in favor of the former test. Spelling efficiency in composition is only very slightly lower than it is in columns. This slight loss is probably due to the distraction of attention by the other factors in writing, such as punctuation, grammatical form, and thought content. The argument that spelling should be taught by using words only in sentences is not very weighty.

(11) **Teaching Homonyms Together or Separately.** Suzzallo and Pearson ('13) report an experiment in which they attempted to determine the relative effectiveness of teaching homonyms together or apart. They used five pairs of homonyms in each of grades 3 to 7 in the Horace Mann school. Each grade was taught in two sections of about equal ability. One section was taught by

the together-method and the other by the separate-method. All other conditions were kept as nearly alike as possible. All classes were tested at the beginning and at the end of the experiment. The outcome showed a decrease of 2.29 errors by the together-method, and of 1.61 errors by the separate-method. There is thus a slight advantage in favor of the former method. Suzzallo's experiment was repeated with the same material by Knight at Montclair, New Jersey, in grades 3 to 7, inclusive. The together-method reduced the errors on the average 2.63% while the separate-method reduced them 2.24%, thus supporting Suzzallo's findings. On the other hand, W. F. Jones ('15) reports that

> "Experiments in teaching homonyms have been made by the department of education at the University of South Dakota, which show that homonyms should not be brought together until the second one of the pair appears in the child's vocabulary. This often gives time to fix the meaning and the spelling of the first member of the pair before the second one appears. In such cases the homonyms give relatively little trouble."

(12) **Class versus Independent Study.** Suzzallo and Pearson further undertook a comparison of progress in spelling when the pupils studied independently, with progress when the pupils studied under supervision. For nearly a year, one class in each of grades 4, 5, and 6 was taught by the supervised-study plan and the other by the independent-study plan. In the latter case, the recitation period consisted chiefly in testing or lesson hearing. Their conclusion is stated thus:

> "The evidence of this experiment, therefore, from whatever angle we study it, shows that teaching of the class-study type is far more effective, than the independent-study type."

TABLE 107. After Suzzallo and Pearson ('13)

Total decrease in errors

Grade	Independent Study						Class Study					
	Room	First Test	Final Test	Net Decrease	Number of Pupils	Average Decrease per Pupil	Room	First Test	Final Test	Net Decrease	Number of Pupils	Average Decrease per Pupil
IV	111	286	228	58	22	2.63 maximum 20	111	275	144	131	22	5.95 maximum 20
V	201	291	208	83	20	4.15 maximum 24	201	291	108	183	20	9.15 maximum 24
VI	206	349	221	128	23	5.56 maximum 32	206	363	143	220	23	9.56 maximum 32
VII	209	351	218	133	20	6.65	209	388	206	182	20	9.10 maximum 32
		1277				Average Decrease per Grade 4.74		1317				Average Decrease per Grade 8.41

(13) **Grouping Words of Similar Spelling.** Is it an advantage to learn words of similar spelling in groups? Wagner [1] made an experiment to answer this question. He divided a 6th grade class into two sections both of which studied the words in the usual manner with the exception that the words were presented to one section in groups, according to their similarity, such as lineal, lineament, linear and lineage while to the other they were presented in miscellaneous combinations. The former section, which studied the words in groups, raised its average percentage from 68.36 in the preliminary test to 97.14 in the final test, or 28.78% while the latter section raised its average percentage from 73.25 to 93.6, or 20.35%. The former group, therefore, gained 8.42% more than the latter, showing a decided advantage in favor of the grouping plan.

(14) **Imagery.** Individuals differ in mental imagery, but recent inquiries point out the probability that pure types of visuals, audiles, motiles, etc., are exceedingly rare and that imagery of any sort may be aroused irrespective of the sense organ through which the stimuli come. Since practically every child possesses images of all classes the safest procedure is no doubt to appeal to a variety

[1] Reported by Pryor. ('15.)

of images. Note here again the discussion of imagery in Chapter XI, p. 166.

(15) **Spelling To-day and Formerly.** Finally it will be of interest to note how the spelling ability of pupils of to-day compares with that of our forefathers, particularly in view of the claim often made that the schools to-day do not train the pupils as thoroughly in the fundamental subjects. One of the comparisons made in the Springfield test (Riley, '08) was that of spelling ability. The same 20 words that had been given as a spelling test to 9th grade pupils in 1846 in Springfield, Massachusetts, were given again in 1906 to 246 pupils of corresponding age in the same school. The pupils in 1846 had made an average grade of 40.6%, while the pupils in 1906 made an average of 51.2%. A similar test, made in Cleveland in the years 1858 and 1909 showed one error less per child in the 1909 test. Apparently the "superior" spelling ability in the good old days is largely an illusion.

CHAPTER XIX

LANGUAGE

Psychological Processes Involved in Language

The subjects thus far considered, namely, reading, writing, and spelling, together with the one to be considered in this chapter, constitute psychologically the complete set of language functions since they all play a part in the communication of ideas. Reading deals with the reception of ideas; while writing, spelling and language in the restricted sense, have to do with the expression of ideas. The term language as used here, and as used in the school program, refers only to that portion of the complete language process which deals with the organization and expression of thought in speaking and writing.

A complete analysis of all the psychological processes involved in the language functions would require again an enumeration of all the elements in reading, writing and spelling. This is unnecessary; and hence our present analysis will be limited to the steps immediately concerned in the expression of ideas in oral or written form. These elements are:

(1) The arising of ideas in the mind.

(2) The simultaneous or successive arousal of symbols or word forms corresponding to the ideas.

(3) The transmission of the nerve impulses, connected with the ideas that arise, to the motor speech centers in oral expression, or to the motor writing centers in written expression.

(4) The transmission of nerve impulses from the latter to either the muscles of the speech organs or to the muscles of the hand and arm.

(5) The execution of the speaking or writing movements.

Steps (3), (4) and (5) are identical with the corresponding ones previously enumerated in the analysis of the writing and spelling processes. The important steps for our present purposes are numbers (1) and (2).

From the psychological standpoint, the arising of ideas in the mind is practically identical in the normal person with the arising of words in the mind, since in the normal child words and meanings are built up together through oft repeated associations of words and meanings. Hence we shall use the term "word-idea" to convey this union of symbol and meaning. Furthermore, from the psychological standpoint, the occurrence in the mind of ideas or words to be expressed, is fundamentally a matter of the psychology of association. Where do the ideas come from? What causes them to arise in the mind? Why do certain ones arise rather than others? To what extent can the occurrence of ideas be controlled? Composition, either oral or written, is simply the outward expression of the ideas that do arise. The occurrence of the ideas and their precise verbal form takes place in the mind as a part of step (1). Obviously then, a study of steps (1) and (2) and an attempt to answer the questions raised thereby, constitute almost entirely a study in the psychology of association or thinking processes. Composing fundamentally is thinking.

Let us take a typical example of oral or written composition, such as a bit of speaking or writing. How do the ideas in this simple composition arise in the mind? The blunt answer is, they arise almost entirely in a mechanical manner according to the established neural or mental connections. The first idea in a chain arises as the result of a stimulus, either through the senses or through previous ideas or images in the mind. No idea probably ever arises independently in the mind as though out of the blue sky, but always in succession to a preceding link, stimulus or occasion. Thinking is simply the flow of ideas, which occurs almost wholly automatically, according to the laws of association —frequency, vividness, primacy and recency. Conversation also is largely the automatic flow of association processes in the minds of the participants. Each statement comes in response to the remark of one of the conversers or in succession to the preceding remarks of the speaker himself. Even more formal composition, such as the writing of a theme, a story or an essay, is principally the result of association processes aroused in the mind by the subject of the theme. This occurs, even in vigorous and original thinking about a subject, largely according to the mechanical laws of association. Voluntary thinking or composing is controlled probably only in two ways: (1) by effort and concentration more ideas are likely to arise than by a purely passive attitude, and (2) by making a selection among the ideas that do arise preference to further chains of ideas are determined. But the arising of ideas themselves occurs fundamentally according to the mechanical, neural and mental connections previously established.

The objection is likely to be made by the reader at this point, that if this is true, how can any new ideas ever arise? or how can any original thinking occur? The blunt answer is: Thinking is original only in the sense of making certain selections, rather than others, among the ideas that do arise and of allowing further associations to arise in connection with them, rather than in connection with others. In this sense, the possibility of original thinking is indefinitely great.

Correct English is fundamentally a matter of associating certain words in certain orders, rather than in others. The reason why people say, "Do it good," is because they have been told since infancy to "Do it good." Language is an immensely intricate net-work of associative links. Wide vocabulary in speaking or writing is due to the arising of numerous words in connection with given meanings. Elegant diction is due to the more appropriate selection of words among the larger varieties that do arise. In the skilled speaker or writer the more appropriate words arise automatically in the course of time. Greater varieties of words come up with given word-ideas because more of them have been previously experienced and retained. Two persons may read the same literary masterpiece, the one may retain a great deal of the actual diction and phraseology in connection with the ideas read, while the other may retain very little. In the course of time the former will acquire a far wider vocabulary and a much nicer diction than the latter because he retains much more of the actual words and forms of expression than the other, and consequently when any topic is brought to his mind, it arouses a far richer wealth of ideas, words and phrases, and by virtue of this wealth he is able to make a far superior selection. Ingersoll appropriately said of Shakespeare:

"The moment his attention was called to any subject—comparisons, definitions, metaphors, and generalizations filled his mind and begged for utterance. His thoughts like bees robbed every blossom in the world, and then with 'merry march' brought the rich booty home 'to the tent royal of their emperor.'" (P. 661, *Modern Eloquence*, Volume V.)

"Some have insisted that Shakespeare must have been a physician, for the reason that he shows such knowledge of medicine, of the symptoms of disease and death; because he was so familiar with the brain, and with insanity in all its forms.

"I do not think he was a physician. He knew too much; his generalizations were too splendid. He had none of the prejudices of that profession in his time. We might as well say that he was a musician, a composer, because we find in 'The Two Gentlemen of Verona' nearly every musical term known in Shakespeare's time.

"Others maintain that he was a lawyer, perfectly acquainted with the forms, with the expressions familiar to that profession. Yet there is nothing to show that he was a lawyer, or that he knew more about law than any intelligent man should know. He was not a lawyer. His sense of justice was never dulled by reading English law.

"Some think he was a botanist, because he named nearly all known plants. Others, that he was an astronomer, a naturalist, because he gave hints and suggestions of nearly all discoveries.

"Some have thought that he must have been a sailor, for the reason that the orders given in the opening of 'The Tempest' were the best that could, under the circumstances, have been given to save the ship.

"For my part, I think there is nothing in the plays to show that he was a lawyer, doctor, botanist, or scientist. He had the observant eyes that really see, the ears that really hear, the brain that retains all pictures, all thoughts, logic as unerring as light, imagination that supplies defects and builds the perfect from a fragment. And these faculties, these aptitudes, working together, account for what he did." (P. 665.)

Thinking and language are the two sides of the same shield. The language used to express ideas depends upon the thinking that goes on in the mind; and the thinking depends upon the verbal-ideational connections established in the neural and mental network by reading and hearing successions of words, phrases and sentences. Language is not words; it is thinking, thinking by means of symbols.

THE MEASUREMENT OF EFFICIENCY IN ENGLISH

(a) *Methods of Measurement.* Four types of measuring devices have been prepared, each for a different phase of language. (1) For measuring the grammatical correctness of language, the writer ('15) has prepared a series of scales consisting of sets of sentences, each stated in two different ways. These sets of sentences are arranged in an order of increasingly difficult steps. The pupil in doing the test is requested to indicate which the correct or preferred form is. A similar scale has been prepared for testing ability in punctuation. (2) For measuring technical knowledge of the grammatical structure of English the writer has devised several tests for ascertaining quickness and accuracy in indicating parts of speech, cases, tenses and modes. (3) For measuring general merit in written composition, two scales have been constructed. The one, known as the Hillegas-Thorndike ('12) scale, is composed of a series of compositions, arranged according to a large number of judgments in the order of steps of increasing merit from zero to nearly one hundred. The second, known as the Harvard-Newton scale (Ballou, '14), is composed of a series of four scales for narration, description, argumentation and exposition respectively. Each scale contains six compositions. These were graded by 24 teachers and approximate in percentage marks the values of 45, 55, 65, 75, 85 and 95. A composition is rated by any one of the scales by comparing its merit with those in the scale, and giving to it the value of the step on the scale to which it is judged equal. (4) Trabue ('16) has prepared a series of language scales, each consisting of some 10 mutilated sentences. These sentences are arranged in steps of increasing difficulty as determined by experimentation with a large number of pupils. In doing the test, the pupil is required to supply the most appropriate words in the blank spaces. It is difficult to say just what this scale measures, but it probably tests the ability to think of the proper word for a given situation.

(b) *Uses and Results.* Without repeating here, we need to say merely that these tests serve the same general purposes for the language functions as the measurements devised for the subjects previously discussed serve for their respective subjects. Measurements of ability in language have shown incredibly large individual differences among the pupils in the same class and the resulting overlapping of the abilities of pupils in succeeding grades. Note the facts in Figure 25 in regard to ability in composition in a certain Illinois high school. The overlapping is enormous and the median gain from year to year is surprisingly small. In fact there is almost no improvement above the first year. The little improvement that is noticeable is probably due chiefly to the dropping out of the poorer pupils in the lower years. Figure 26 shows similar facts with regard to ability in recognizing correct grammatical forms.

Brown and Haggerty ('17) measured the progress in composition ability in the case of 78 pupils in three classes during a period of 12 weeks by having them write an extemporaneous composition each week. These compositions were rated by the Harvard-Newton scale. The median progress of the three classes is indicated in

Figure 74, being 4.2 points for Classes I and III, and 5.2 points for Class II. These graphs indicate a noticeable improvement

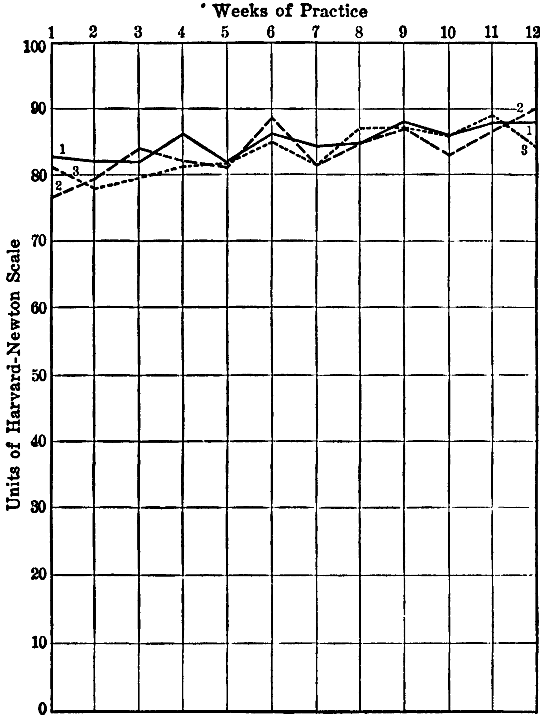

FIG. 74.—Median scores in composition for three high-school classes. Continuous line represents a first-semester freshman class; the dotted line represents a second-semester freshman class; and the dash-line represents a first-semester sophomore class. After Brown and Haggerty ('17, p. 524).

from the first week to the twelfth, but show no difference in ability between the sophomore class and the freshman classes.

ECONOMIC METHODS FOR ACQUIRING SKILL IN THE ENGLISH LANGUAGE

A great variety of experiences has been obtained and an equally great variety of personal opinions is held by teachers with regard to the most effective manner of learning and teaching English. But there is a corresponding paucity of scientific facts regarding the matter so that we cannot speak with certainty concerning very many of the factors and conditions that promote or retard development of skill in the use of language.

(1) **The Acquisition of Ideas.** According to the psychological viewpoint, from which we are here examining the language process, the first and most fundamental element is the acquisition of ideas. How this may be done our present psychological knowledge gives no specific directions other than the common-sense advice of gaining as great a variety of ideas as possible through direct personal experiences and contact with environment and through extensive reading. The extent to which such varied experiences and wide reading will be profitable depends mainly upon the extent to which the experiences and facts are assimilated and retained. Next to the experiences and ideas themselves, excellent retentive capacity, or memory is a very essential factor. Re-thinking the experiences and facts is likely to be a most helpful exercise in making new ideas a part of one's mental machinery.

(2) **The Acquisition of Words and Forms of Expression.** The next important element is the acquisition of extensive vocabulary and phraseology. On this point also little more can be said at present than what, on the basis of general experience rather than on the basis of demonstrated facts, common sense and good judgment dictate. Wide vocabulary and precise phraseology can probably be acquired mainly (a) by reading and hearing language which employs extensive vocabulary and proper phraseology, and (b) through definite attempts to express one's thought by means of these words and phrases. Probably one of the best means of extending one's vocabulary is to take special account of new words as one reads or hears them, and to make a special point of using them on the first, and on every other, appropriate occasion. A useful plan is to record these new and unfamiliar words in a special notebook. They must, however, not be kept in the notebook only, but they must be referred to frequently so as to make them a part of the mental fabric of connections to be used in thinking, speaking and writing. Pupils often are required to keep such notebooks, but usually the words never get farther than the notebook. Special effort must be made to use the words so that they may gradually be woven into the automatic association processes of the mind.

Another plan is to make a list of the less familiar or unknown words that are observed in the language of good writers, speakers and conversationalists, and to think of as many synonyms and equivalent phrases as possible so that when occasion for using one of the words arises, synonymous terms will come up which might preferably be used instead. Frequent drills and tests might very profitably be given in accordance with this suggestion. The teacher could give such a list of words in mimeographed form to the pupils and then see how many corresponding synonyms in a given time limit, say 5 or 10 minutes, each pupil can write out. The habitual use of words is fundamentally a matter of association processes and these can be built up only by a repeated use of the words to be acquired.

For such instruction it would be exceedingly useful to have a list of the 10,000 most common words employed by the best current writers. Special drill exercises in the use and meaning of these words could be devised by the teacher.[1]

(c) The influence of knowledge of other languages. It has been generally urged that a knowledge of other languages, particularly of those from which many English words have been derived, is of much value in extending the vocabulary, in developing a finer discrimination in the meanings and uses of words, and in facilitating the writing of English. The objective data thus far available on this problem have been summarized in the chapter on transference of training, to which the reader should turn for a reconsideration of that evidence. In general the evidence indicates (1) that the study of Latin seems to increase the size of a pupil's English vocabulary only slightly when Latin is taught as it ordinarily has been taught, or quite considerably when taught with specific reference to the derivation of English words; and (2) that, when the differences in the original abilities of pupils are allowed for, the study of other languages aids either very little or not at all in the writing of English.

(3) **Acquisition of Grammatically Correct English.** (a) The influence of knowledge of grammar. Grammar was introduced into

[1] Such a list has been prepared by the writer on the basis of the vocabulary studies described in Chapter XVIII. The Starch List of 5,903 different words used by 40 authors was checked up with the Eldridge, Ayres, Jones, and Cook Lists to find all additional words in them which were not in the Starch List. This made a total of 9079 different words found in 40,000 running words (Starch List) of 40 different high-grade current magazine writers, in some 43,000 running words (Eldridge List) of newspaper writing, in some 23,000 running words (Ayres List) of correspondence, in 15,000,000 running words (Jones List) of children's compositions, and in 200,000 running words (Cook List) of family correspondence.

the schools as a scientific study of the structure of the language with the belief that this knowledge would aid in the acquisition of correct usage of the language. To what extent is this belief justified? All definite investigations of this problem tend to show that knowledge of technical grammar is of much less service in developing use of correct English than has always been supposed. This experimental evidence was reviewed in Chapter XIV on the transference of training, to which the reader should turn, and hence it will not be restated here. Furthermore, the studies by Charters ('15) on grammatical errors in oral and written language indicate that a relatively small number of grammatical rules is concerned in ordinary language. In accord with these results, the amount of time devoted to the study of grammar has been materially reduced in recent years and a great many of the topics have been entirely eliminated. The amount of grammatical knowledge which functions directly in establishing correct usage is relatively small. For example, the Iowa Committee on Minimum Essentials recommended in 1915 the elimination of the following topics:

"The exclamatory sentence; the interjection; the appositive; the nominative of address; the nominative of exclamation; the objective complement; the adverbial objective; the indefinite pronouns; the objective used as a substantive; the classification of adverbs; the noun clause; conjunctive adverbs; the retained objective; the modes (except possibly the subjunctive of 'to be'); the infinitive; the objective subject; the participle, except the definition and the present and past forms; the nominative absolute; the gerund nominative absolute; sentences for analysis and parsing that involve subtle points of grammar; formal parsing; conjugation; diagramming; person of nouns." (After Betts, '17.)

(b) The function of imitation. These investigations imply that linguistic forms and expressions are acquired very largely through imitation, both conscious and unconscious, perhaps especially the latter, of the forms and expressions read or heard, particularly those heard in one's customary environment. Language forms are psychological habits which become deeply ingrained in the human psycho-physical system through constant repetition. To hear and say, from birth on, "good" for "well" in such expressions as, "Do it good," or "I don't feel good," establishes such strong chains of associated bonds that, in spite of better knowledge, such expressions will continue to be used and not be overcome by pages of grammatical knowledge.

The factor of conscious, intentional imitation is probably employed far too little in the development of ability in composition. Students of art learn to paint by copying the great masterpieces. Students of composition are shown great masterpieces and asked to read them but are told not to imitate them as that would prevent the development of originality. Is this not incorrect procedure, psychologically? Why would it not be excellent practice in acquiring vocabulary, diction, style, and even ideas to have pupils read and study, for example, a description of a landscape and then to ask them to write one of a real landscape in their environment by making it just as similar to the masterpiece, in plan, diction and phraseology, as the actual landscape permits. Or why should not such a plan as the following be really effective in acquiring the use of words. Have the teacher take such a description as the following one of George Washington:

"When Washington was elected general of the army he was forty-three years of age. In stature he a little exceeded six feet; his limbs were sinewy and well proportioned; his chest broad, his figure stately, blending dignity of presence with ease of manner. His robust constitution had been tried and invigorated by his early life in the wilderness, his habit of occupation out of doors, and his rigid temperance, so that few equalled him in strength of arm or power of endurance. His complexion was florid, his hair dark brown, his head in shape perfectly round. His broad nostrils seemed formed to give expression and escape to scornful anger. His dark blue eyes, which were deeply set, had an expression of resignation and an earnestness that was almost sad."

Let the teacher select all the important descriptive words and phrases from it and put them in a list, such as this:

stature
sinewy
well-proportioned
stately
blending dignity
ease of manner
robust
invigorated
occupation
rigid temperance
endurance
complexion
florid
nostrils
scornful
resignation
earnestness

The pupils should first study these words and phrases to make sure that they thoroughly understand them; then they should write a description of some person of their acquaintance who may also be known to the teacher, using as many of these phrases as may be applicable to this person. The pupils in this case should not first read or hear the description from which these words were taken. After they have written their own composition, they should then compare it with the masterpiece so that they could check up their own use of words with the possibly better use in the model.

Such a plan would seem to be worth while experimenting with to ascertain its proficiency in developing language ability. How can a pupil acquire proper words to be used in describing a human being if he does not know, and has no means of finding out, what such words are? Such a plan would seem to have on a *priori* grounds two distinct advantages: First, it would make conscious use of imitation which is unquestionably a potent factor in the acquisition of language; and second, it would tend to make the learning of language definite and concrete. The great difficulty in learning composition is that the advice and suggestions given by the teacher are too vague and indefinite. The pupil is told that he lacks organization, that he must develop a better vocabulary, or that he lacks imagination. But how is any child specifically and concretely to know how to improve in these respects? How may a child know that such words and phrases as "stature," "stately dignity," "florid complexion" might be the most appropriate to use when describing certain persons? When the child learns addition or spelling he knows more precisely what he has to learn. Instruction in composition can possibly not be particularized as fully, at least not as easily, but that should not prevent as much particularization as possible.

The chief objection by teachers of English to such a procedure is that they believe it would kill originality and make mere thoughtless, verbal machines out of their pupils. In answer to this point, however, we must remember that the most original people in the world are also the ones who use to the fullest extent the work, methods, and ideas of others. The most original persons are also the most imitative persons. Ingersoll said of Shakespeare:

"Of course Shakespeare made use of the work of others, and we might almost say, of all others. Every writer must use the work of others. The only question is, how the accomplishments of other minds are used, whether as a foundation to build higher, or whether stolen to the end that the thief may make a reputation for himself, without adding to the great structure of literature.

"Thousands of people have stolen stones from the Coliseum to make huts for themselves. Thousands of writers have taken the thoughts of others with which to adorn themselves. These are plagiarists. But the man who takes the thought of another, adds to it, gives it intensity and poetic form, throb and life, is in the highest sense original.

"Shakespeare found nearly all of his facts in the writings of others and was indebted to others for most of the stories of his plays. The question is not: Who furnished the stone, or who owned the quarry, but who chiseled the statue?" (P. 645, *Modern Eloquence.*)

Some of the prominent literary writers have not only pointed out the importance of conscious imitation in the development of their own styles of writing, but have described their own conscious attempts to imitate other great writers. For example, Stevenson has said this concerning imitation:

"That, like it or not, is the way to learn to write. It was so Keats learned, and there never was finer temperament for literature than Keats's; it is so, if we could trace it out, that all men have learned. Perhaps I hear some one cry out: 'But that is not the way to be original!' It is not, nor is there any way but to be born so. Nor yet, if you are born original, is there anything in this training that shall clip the wings of your originality. There can be no one more original than Montaigne, neither could any be more unlike than Cicery; yet no craftsman can fail to see how much the one in this time tried to imitate the other. Burns is the very type of a prime force in letters; he was of all men the most imitative. Shakespeare himself, the imperial, proceeds directly from a school. Nor is there anything here that could astonish the considerate. Before he can tell what cadences he truly prefers, the student should have tried all that are possible; before he can choose and preserve a fitting key of words, he should long have practiced the literary scales . . . and it is the great point of these imitations, that there still shines beyond the student's reach his imitable model."

"Whenever I read a book or passage that particularly pleased me, I must sit down at once and set myself to imitate that quality of propriety or conspicuous force or happy distinction in style. I was unsuccessful and I knew it, but I got some practice in these vain bouts in rhythm, in harmony, in construction, and in co-ordination of parts. I have thus played the sedulous ape to Hazlitt, to Lamb, to Wordsworth, to Browne, to Defoe, to Hawthorne, to Montaigne, to Baudelaire, and to Obermann." (Stevenson, R. L., *Memories and Portraits*, p. 55.)

Franklin likewise described his attempt to improve his own style of writing by carefully studying a volume of *The Spectator* and developing a style similar to it.

"I read it over and over and was much delighted with it. I thought the writing was excellent, and wished, if possible, to imitate it. With that view I took some of the papers and making short hints of the sentiments in each sentence, laid them by a few days, and then, without looking at the work, tried to complete the papers again, expressing each hinted sentiment at length, and as fully as it had been expressed before, in any suitable words that should occur to me. Then I compared my 'Spectator' with the original, discovered some of my faults, and corrected them." To acquire a stock of words and a readiness in recollection and use of them, he "took some of the tales in the 'Spectator' and turned them into verse; and after a time when I had pretty well forgotten the prose, turned them back again." (Quoted by Bolton, F. E., *Principles of Education*, p. 421.)

Brander Matthews in a more general way has pointed out the relative shares of imitation and originality.

"Consciously or unconsciously every artist is a debtor to the past. The most original of innovators has made his originality partly out of himself, partly out of what he has appropriated and absorbed from those who practiced his art before him. Only a few of his separate contrivances are his own, and the most he can claim is a patent on the combination.

". . . The young artist is a weakling if he openly robs any single one of his predecessors; he is a dolt if he does not borrow from as many of them as may have the separate qualities he is striving to combine.

"The arts are one in reality; and what is true of painting and sculpture and architecture is true also of literature, of prose and verse. For example, there are few men of letters of our time whose prose has been more praised for its freshness and its individuality than the late Robert Louis Stevenson; but his was an originality compounded of many samples. He confessed frankly that he had sat at the feet of the masters, playing the 'sedulous ape' to a dozen or more, and at last slowly learning how to be himself. Again, the verse of Dante Gabriel Rossetti has a note of its own, a note which many younger poets have delighted to echo and re-echo; but Rossetti told a friend that the exciting cause of his 'Blessed Damozel' was the 'Raven' of Edgar Allan Poe, and Poe's own indebtedness to Coleridge is obvious, even if it had not been expressly avowed."

(c) Specific attention to frequently recurring grammatical errors. Specific attention to, and special drill in, the precise functions to be developed have been shown over and over again by experiment and experience to be most directly efficacious in producing improvement. Several valuable studies have been made in recent years of the most common types of errors in oral and written language and of the frequency of their occurrence. Charters and Miller ('15) made a study in coöperation with the teachers in schools in Kansas City, by keeping a record for one week of all errors in oral language heard by the teachers in and about the school, and by collecting all errors in the written work of the pupils for one month. A similar study was made by H. D. Fillers ('17), at Bonham, Texas, based on results from 900 pupils in grades three to eight for written language, and in grades two to eight for oral language.

A third study has been reported by Sears and Diebel ('16) from Cincinnati based on oral errors of 1,378 children of grades three to eight which were heard by the teachers during a period of four days.

A fourth investigation was made by Starch (unpublished) in which oral errors of pupils in grades one to eight were collected by the teachers, and oral errors of university students by a group of special students, and written errors were collected from some 1,700 themes obtained from eleven high schools. The high schools varied from very small ones to one having more than 1,000 students. The total number of oral errors collected was 2,916 from grade pupils and 1,164 from university students. The total number of written errors collected from the high school pupils was 2,316, making a total of 6,396 errors. The results of all four studies are summarized in parallel columns in Table 108.

On the whole, the corresponding parts of all four studies agree surprisingly well, showing considerable reliability as well as striking similarity in the types of errors found in different parts of the country.

TABLE 108
Errors in language

Nature of Error	Written					Oral				
	Charters		Fillers		Starch	Charters	Fillers	Sears & Diebel	Starch	
	% of Total	% of First 21 Rubrics	% of Total	% of First 21 Rubrics	% of Total	% of Total	% of Total	% of Total	% of Total	
									Elem. Sch'ls	University
1. Subject of verb not in nominative case. Ex. Me and some boys went..... ..	0+	1	0	0	0.2	4	5		4 5	1 1
2 Predicate nominative not in nominative case. Ex. It was not him..	0+	1	0	0	0.1	2	0	2	0 7	3 9
3. Object of verb or preposition not in objective case. Ex. She appointed Lynawood and I....	0+	1	0	0	0 3	1	1	0+	0 3	5 9
4. Wrong form of noun or pronoun. Ex. Theirself, hisself	5	16	7	11	3 9	2	1	0+	1 4	0 7
5. First personal pronoun standing first in a series. Ex. Me and him went	0+	1	1	1	0 1	2	2	5	4 0	0 0
6. Failure of pronoun to agree with its noun in number and gender Ex. Nobody can do what they like......	1	4	1	1	6 1	0+	0	0+	0 1	4 1
7. Confusion of demonstrative adjective and personal pronoun. Ex. Them weeds, them things, them girls.....	0+	0+	0	0	0 0	3	4	4	1 7	02.
8. Failure of verb to agree with its subject in person and number. Ex. They was brought	6	19	1	7	4 6	14	9	6	10 4	11.4
9. Confusion of past tense and past participle. Ex. I seen him......	2	5	2	4	1 2	21	20	19	13 3	5.9
10. Confusion of past and present tense. Ex. They come along and took mine......	4	12	12	19	0 4	2	5	3	6.0	1.8
11. Wrong tense form Ex. John's dog went home..............	2	5	2	4	0 5	5	3	15	3.7	0.9
12. Wrong verb. Ex. Can we stay? ...	2	7	2	4	14 6	12	21	14	18.6	15.5
13. Incorrect use of word. Ex. If you was to be tardy.............	0−	2	0	0	5.1	0+	0	0+	0.0	2.3
14. Incorrect comparison of adjectives. Ex. The winds are much more cooler ..	0+	0+	0	0	0.1	1	0	1	0 6	
15. Confusion of comparative and superlatives. Ex. Jolliest (of two)..	2	6	0	0		0+	0	0+		

TABLE 108—Continued

Nature of Error	Written					Oral				
	Charters		Fillers		Starch	Charters	Fillers	Sears & Diebel	Starch	
	% of Total	% of First 21 Rubrics	% of Total	% of First 21 Rubrics	% of Total	% of Total	% of Total	% of Total	% of Total: Elem. Sch'ls	% of Total: University
16. Confusion of adjective and adverb Ex. Will that there do?	0+	0+	3	6	2.8	4	2	2	0.9	4.1
17. Misplaced modifier. Ex I only went once.	2	6	2	4	2.8	0			0.6	3.2
18. Double negative. Ex. I don't have no pencil	0+	1	0	0	1.1	11	14	12	10.3	4.1
19 Confusion of preposition and conjunction. Ex. He looks like he is rich	0+	1	1	1	3.6	0		1		
20. Syntactical redundance. Ex. Where's he at?	4	11	5	7	3.3	10	10	11	16.6	9.6
21. Wrong part of speech due to similarity of sound. Ex. I would be known	11		16	25	3.1	1	0	0+	0.0	0.8
22. Failure to put period at end of sentence	30		13							
23. Failure to put question mark at end of question	2		0							
24. Failure to put apostrophe to denote possession	6		2							
25. Omission of subject	3		0							
26. Omission of predicate	2									
27. Confusion of dependent and independent clause	9		8							
28. Failure to begin sentence with capital letter			8							
29. Use of wrong article										
30. Failure to use quotation marks										
31. Wrong word or phrase. *That* for *The*					10.3				2.5	5.8
32. Words omitted. I been to —					10.3				0.3	1.7
33. Miscellaneous					7.5				3.2	

These studies show that most of the errors in language are confined to a small number of types. Thus in Charter's tabulation, 71% of all oral errors fall under five types or grammatical rules, namely, No. 9, Confusion of Past Tense and Past Participle, 24%;

No. 8, Failure of Verb to Agree with its Subject in Number and Person, 14%; No. 12, Wrong verb, 12%; No. 18, Double Negative, 11%; and No. 20, Syntactical Redundance, 10%. The same situation obtains among written errors; 91% of all errors fall under ten classes or rules.

There are several rather striking differences between the errors of written and oral composition. Written composition is much more apt to get the wrong form of noun or pronoun (16%, 11%) than oral (2%, 1%, 0 – %) and the same is true of confusion of the tenses (12%, 19%, vs. 2%, 5%, 3%). On the other hand, oral language is much more apt to confuse the past tense and past participle (24%, 20%, 19%, vs. 5%, 4%), to use the wrong verb (12%, 21%, 14%, vs. 7%, 4%), and to use the double negative (11%, 14%, 12% vs. 1%, 0%).

Sears and Diebel also tabulated their material to show the relative proportions of different types of errors for the various grades. (Table 109.)

TABLE 109

Classification of Errors	Grades 3	4	5	6	7	8	Total
1. Verbs	44.2	60.0	55.4	54 9	43 3	48 2	49 9
2. Pronouns	15.9	14.0	6.7	7.7	12.3	18.8	13.5
3. Negatives	11.5	7.1	20 2	7.3	15.2	14.0	11.6
4. Syntactical redundance	8.0	6.6	11.2	12.6	16.5	9.6	9.7
5. Mispronunciations	14 7	7.8	2 2	4.9		1.7	8.0
6. Prepositions	3.4	3.2	1.8	5.6	4.1	2.6	3.5
7. Adjectives and adverbs	2.0	0.6	2.2	6.6	8.2	4.8	3.3
8. Ambiguous expressions		0.2					.2

Percentages of errors in each grade due to each class of mistakes. The forty-one specific errors which Sears and Diebel found most frequent follow with their respective frequencies:—

1. haven't no for haven't any 233
2. seen—had saw 180
3. ain't for am not, isn't, aren't 124
4. done 113
5. got, ain't got, haven't got 112
6. I and my brother 96
7. kin, jist, git, kitch 91
8. ain't for haven't, hasn't 89
9. Frank and me for Frank and I 80
10. is for are 76
11. them for those 75
12. learn for teach 71

13. can for may	60
14. my mother, she	58
15. got for receive, become, grow, is	53
16. that there	38
17. don't for doesn't	36
18. It was me	36
19. leave for let	34
20. went for gone	32
21. come for came	31
22. never gave	30
23. by my aunt's	28
24. drawed, throwed, growed, knowed	27
25. somepin, for something	25
26. broke for broken	22
27. lay for lie	21
28. make dinner for prepare, get	21
29. says for said	20
30. all two, all both	19
31. readin, nothin	18
32. by us for near us	16
33. he does it like she does	15
34. why, and, so, at the beginning of sentence or in middle of sentence	15
35. that, which, for who and whose	14
36. onct	12
37. in back of	12
38. funny, lots, etc., for queer, many	11
39. et for ate	11
40. run for ran	11
41. set for sit	10

R. I. Johnson obtained samples of written exposition, narration, and description from 132 high school freshmen and 66 college freshmen from the Kansas City High School and Junior College in an effort to determine the persistence of errors in written English. The high-school freshmen made 2,160 errors in 50,371 words, while the college freshmen made 787 errors in 32,693 words. Roughly this was 23 errors per thousand words for the former and 42 errors per thousand words for the latter. The college freshmen are therefore distinctly superior. The following table shows the relative improvement for the various types of errors. Naturally the raw results of such an experiment somewhat exaggerated the improvement due to four years of training because of the elimination of the poorer high school students before reaching college. In the last column of the table this factor has been compensated for.

TABLE 110

Types of Errors	Total No. Errors of 132 High School Freshmen	Errors of 66 College Freshmen Increased to Proportion of 132	Per Cent Decrease of Errors of College Freshmen	Rank in Prevalence	Rank in Persistency	Per Cent Decrease in Errors of College Freshmen after Elimination has been Allowed for
1 Mistake in case of pronoun ...	11	3 0	72.7	14	13	71
2 Other errors with pronoun	102	61 6	40 0	6	7	36
3. Use of verb	93	49 0	47 0	7	8	37
4. Adjective and adverb	52	43 0	17 3	9	1	-11 (increase)
5. Prepositions and conjunctions .	50	37.0	26.0	10	5	21
6. Ungrammatical sentence structure	220	49 0	77.7	8	11	74
7. Unclear meaning	46	26 0	43.5	12	4	17
8 Mistakes in punctuation	232	154 8	33.0	3	6	28.3
9. Use of apostrophe............	150	115.5	23 0	5	3	16 5
10. Capitalization...............	196	184.8	5.7	2	2	-6.5 (increase)
11. Careless omission or repetition..	223	118.6	47 0	4	9	43.
12. Mistakes in spelling...........	676	320 3	52 5	1	11	47 4
13 Quotation marks	25	13 8	44 8	13	10	46.3
14. Miscellaneous errors	85	35 4	58 3	11	12	54

I have computed the correlation of the two series of ranks in the above table by the Spearman rank method which yields .26. This means that there is slight connection between the prevalence and the persistency of errors, that is, if there is any tendency at all it is for the more common errors to be also more persistent.

Such studies should prove exceedingly useful in helping the teacher to devise special exercises and drills designed to eliminate these errors and to substitute for them correct forms of expression.

(4) **Oral versus Written Practice in Composition.** The processes of oral and written composition differ psychologically from each other in certain important respects; and the same is even more true of formal oral composition and everyday speech. Ordinary conversation consists for the most part of short verbal responses each lasting but a few seconds, initiated by continuously recurring stimulations from the other individuals engaged. Formal oral composition on the other hand is a reaction many times as long with but a single formal impulse at the beginning. With this enormously increased length of reaction arise all the complex problems of structure and the accompanying strain on the attention, to say nothing of the powerful emotional inhibitions of thought and action resulting from self-consciousness and fear of failure, which practically do not exist at all in ordinary spontaneous speech.

Again, written composition differs psychologically very much from oral composition in permitting much more deliberate action. If a satisfactory thought does not present itself, there is no imperative necessity for immediate action at all costs. There are no distractions of the attention from a staring audience. Perhaps most different of all, written composition presents an opportunity for revision of structure, choice of material and words which oral composition cannot permit. Even the vocabulary and sentence structure is different to a certain extent.

It has already been pointed out in the chapter on transfer of training that psychological reactions tend strongly to be limited closely to the conditions under which training takes place. The moral as to composition is obvious. Oral composition no less than written composition requires specific practice.

Probably over nine-tenths of the composition of the average adult is oral and only a very small fraction is written. And yet the schools until recently have directed about nine-tenths of all specific training in composition toward written expression, either with the belief that training in writing would carry over directly to oral composition or with no realization at all of the importance of oral composition. The whole process has disregarded the psychological law of association that associated bonds should be formed in the order and manner in which they are to be used. If language is to be spoken, the neural links via the motor speech centers should be specifically and correctly exercised. It is needless to point out here the inestimable value of speaking and conversing in correct, convincing manner. Why should the schools not direct their efforts far more fully to specific training in oral composition? The reply might be made that all speaking in general and all reciting in classes affords practice in oral composition. But the trouble is that most of it is done with little or no attention to correct, well-organized expression. Gross errors, to be sure, are pointed out to the child; but why should the school not assign a theme to each pupil in the English class on which he is to give a 3 or 4 minute oral composition in correct, well-planned and thought-out form? Apropos of this point, a committee of The Illinois Association of Teachers of English arranged

"A course for the second semester of the ninth grade, which was to be taught in two ways; one class would have only written exercises; the other, a combination of two-thirds oral and one-third written. All classes taking either course were to be given the same written tests at the beginning, at the middle, and at the end of the semester. All the papers written by each class, including these tests, were to be forwarded to the committee in charge of the experiment, accompanied by a report from the teacher, stating as accurately as possible how much time he spent in preparation, in conference, and in correcting papers, and also his opinion as to the results of the experiment.

"The outcome was decidedly favorable to the use of oral composition. The sections taking the combined course were better at the end of the semester in thought-vigor, freedom and interest than the others; they were no worse in spelling and punctuation and better in handwriting—indeed, the writing sections showed marked degeneration in all matters of mechanics. Over half of the 22 schools which carried out the experiment in full reported greater improvement in the combination sections, while only two reported less improvement." (Reported by Hosic, '15, pp. 98–99.)

(5) **How Much Written Composition is Profitable?** Written composition, even aside from consuming the time that might more profitably be used for oral composition, is probably overdone in many English classes. So much is required that it becomes a bore with little profit to be derived for the amount of time consumed in putting down upon paper a few trivial ideas about fictitious and worthless subjects. Professor Lounsbury and other teachers of English have begun to doubt the value of such extensive writing. He says:

"I am by no means disposed to go so far as the historian of New England, John Gorham Palfrey, who as I have been told, was wont to express the desire that an act of Congress should be passed forbidding on pain of death anyone under twenty-one years of age to write a sentence. Excess in one direction can not be remedied by excess in the opposite. Still, none the less am I thoroughly convinced that altogether undue importance is attached to exercises in English composition, especially compulsory exercises; that the benefits to be derived from the general practice in schools is vastly overrated; that the criticism of themes, even when it is fully competent, is in the majority of cases of little value to the recipient; that in a large number of instances the criticism is and must ever be more or less incompetent; and that when the corrections which are made inefficiently and unintelligently, as is too often the case, the results are distinctly more harmful than helpful."

William Lyon Phelps, Professor of English Literature at Yale, has reached a similar conclusion:

"On the subject of required English composition, I am a stout, unabashed and thorough sceptic. And although the majority is still against me, I am in good company. Professor Child read and corrected themes at Harvard for about forty years: at the end of the time it was his fervent belief that not only was the work unprofitable to the student, but that in many cases it was injurious. That it is always injurious to the instructor, when it is intemperately indulged, is certain. When I was an instructor at Harvard, I one day met Professor Child in the yard. He stopped a moment and asked me what kind of work I was doing. I said, 'Reading themes.' He put his hand affectionately on my shoulder, and remarked with that wonderful smile of his, in which kindness was mingled with the regret of forty years, 'Don't spoil your youth.' Professor Wendell, who inherited the bondage under which his predecessor groaned, has never really believed in the efficacy of the work. Professor Lounsbury of Yale has given valuable and powerful testimony against it. Professor Cook and Professor Beers—two quite different types of men—are in this point in absolute agreement." (P. 117.)

By way of concrete evidence for his opinion, Professor Phelps tells that after requiring only a moderate amount of theme writing during the freshman and sophomore years he submitted a batch of compositions written by his juniors to one of the Harvard professors. He read them carefully and testified that they were exactly as good technically as those done by Harvard juniors. He further says:

"I know of nothing in the world that illustrates more beautifully the law of diminishing returns than required courses in composition. A class of students will never under any circumstances write five times as well by writing five themes as they will by writing one; but the reading and correcting of five themes require five times the effort on the part of the body of teachers." (P. 123.)

It is evident, from the diversity of opinion and practice regarding the question, that a careful and extensive investigation should be made to determine if possible the optimum amount of writing to be required in English courses. The report of the Committee of the Modern Language Association of America and the National Council of Teachers of English on the Cost and Labor of English Teaching (1913), based on returns from 552 English teachers in 93 high schools and 345 English teachers in 96 colleges, states that

"The amount actually written under present conditions averages for high schools 380 words a week throughout the year, and for colleges 630 words a week. Ideal conditions would slightly increase these averages to about 430 for high schools and 680 for colleges, and would make possible equal attention to oral and to written training."

These facts are interesting as indicative of the present practice; they are significant as indicative of the consensus of opinions concerning the ideal amount of writing to be done. No one of course knows what the ideal amount is nor where the region of diminishing returns is located. Opinions do not settle the question. Experimental and statistical inquiries must be made to determine the actual results produced by various amounts of writing under varying conditions and to locate, if possible, the region of diminishing returns.

(6) Good English in All Oral and Written Work. One of the great counteracting forces to the influence of the teaching of composition in English classes, particularly in high schools and universities, is the utter disregard for proper use of language in work outside of the classes in English. Pupils wear their Sunday clothes in only one class and go shabby the rest of the time. Instruction in composition is effective only when it succeeds in making correct, elegant language a part of the automatic association processes of thinking. During nine-tenths of the time, the language habits are being mechanized in a slipshod manner. How can the instruction in English classes prevail against the iron chains of habit formed during the rest of the day?

In his classic chapter on habit, James says: "The second maxim is: Never suffer an exception to occur till the new habit is securely rooted in your life. Each lapse is like the fall of a ball of string which one is carefully winding up; a single slip undoes more than a great many turns will wind again."

Pupils should be required to be as careful in all their oral and written work as they are in their classes in English. The organization of schools should be so modified that the necessary coöperation between teachers of English and teachers of other subjects could be made possible. Why should not all written work in all other subjects have to pass muster before the teacher of English? While all teachers, besides teachers of English, are in theory supposed to watch carefully over the English used by their pupils, they do as a matter of fact pay very little attention to it, partly because they feel they are too busy teaching the content of their own subjects to be able to spare the time, and partly because many of them are not as competent to correct the English of their pupils as they should be.

A powerful incentive to the pupil to use at all times the best language that he is capable of could be given by basing his English mark to the extent of say one-third or one-fourth, upon his work in the English class, and to the extent of two-thirds or three-fourths upon the quality and correctness of his English in other classes.

(7) **Practice in Expressing Really Important and Personally Vital Ideas.** On the assumption that thought and language are intimately related, that in fact for practical purposes idea and word or thought and language are substantially identical. It would seem to be a fair inference that the exercise in oral and written composition should be carried on in conjunction with topics concerning which the pupil really has ideas or concerning which, if he has none, it would be really worth while to acquire ideas. From a psychological and humanitarian standpoint, it seems almost a crime to ask a pupil to write several pages of words on a topic concerning which he has no ideas or in which he has not the slightest interest, or which has no vital importance to him or to anybody else. Yet how often is such the case! It is not likely to be very stimulating to a pupil to be asked to write 150 words with this phrase as the beginning of the first sentence: "Life to me seems . . . . "; to write a theme on that veteran of topics "The Autobiography"; or on such subjects as "Why Go to Church?" "Why go to College?" "What Picture Impressed me Most," "What I got out of Virgil's Æneid."

Brown and Haggerty incidentally point out in their study that the pupils wrote distinctly better compositions on certain topics than on others. The most striking case was the topic of the sixth week, "How I Earned Some Money" which produced a rise in the curves to as high a point as was reached even at the end of the twelve weeks. The topic for the fifth week, "The Pleasures of Skating (or some other sport)" and for the seventh week, "The Right Kind of a Chum" produced distinct drops in the curves.

Too much of the practice in theme writing has to do with fictitious situations and not with vital ideas. The average adult, unless he enters a profession of which literary work is a necessary part, rarely has occasion to write about fictitious or far-fetched subjects. His composition, both oral and written, has to do with problems which are vital to his life and concerning which he actually has ideas or concerning which, if he has none, it is very important for him to acquire ideas. His welfare, his business and professional success may depend upon the forcefulness of his letter or upon the convincingness of his interview. Many adults testify to the effect that they obtained most of their training in organizing and expressing their thoughts in connection with their business, professional, and vocational problems.

Why should not the composition in school be centered around problems that are real to the pupil or that will be real to him? Why should not the written work in other school subjects be submitted as the work in English composition, and why should not the correction and development of language be centered around the expression of thoughts actually necessary in his school work? The engineering school in a western university severed its instruction in English from the academic English department because of the purely formal drill that was provided for its students, and instead secured the services of a teacher, who was trained both in engineering and in English, to give instruction in English by using their writing, their laboratory notes and reports, their language work in the various courses as the chief basis of instruction in English.

(8) Effect of Differences in Teaching Ability. The differences of the ability of teachers are probably as striking, if not more so, in the teaching of English as in any other branch. The differences in the average abilities of entire classes are very great. The following table gives the median value of the compositions written on the same topic and under similar circumstances in the high schools in a certain county in Illinois. Each composition was rated by three or more judges according to the Hillegas-Thorndike scale.

TABLE 111

High School	Number of Pupils	Median Score in Composition (Hillegas Scale)
1	6	40.5
2	77	47.4
3	169	47.7
4	15	52.8
5	41	52.9
6	113	58.5
7	...	62.3
8	55	63.0
9	261	63.1
10	22	77.2

From this table it appears that the pupils in the best school wrote compositions of approximately twice as good quality as the pupils in the worst school. Furthermore there is, so far as these limited data go, no indication of connection between size of school and quality of composition. The probability is that the most potent factor in the situation was the teaching ability of the instructor.

CHAPTER XX

ARITHMETIC

The Psychological Processes Involved in Arithmetical Operations

An analysis of the psychological processes involved in arithmetical operations may be undertaken in two ways: either we may analyze the mental activities concerned in a typical arithmetical operation as performed by an adult or by a practiced pupil, or we may trace the genetic development and combination of arithmetical concepts and processes in the child. The one would be a dissection of the finished product as carried out by the trained individual; the other would be a synthesis of the arithmetical elements as they arise in the mental growth of the child. Both methods of attack will lead essentially to the same set of elements. Let us pursue for the present the first method of approach and let us take as an illustration a common everyday type of arithmetical computation. Suppose you purchase at a store, fruit for 10c., a loaf of bread for 10c., a pound of butter at 45c., and a bunch of celery for 7c., and give the clerk a dollar bill, how much change should you receive? What mental processes are either involved or presupposed in arriving at the answer, 28c., or rather in the clerk's making the change up from 72 until he reaches one dollar? An analysis will show at least the following steps: (1) The concepts of numbers and their meaning, (2) the ability (a) to hear, interpret and speak (if only by inner speech, which accompanies even the thinking of the numbers) the sounds for the numbers when the calculation is carried out mentally, or (b) to write and read the symbols for the numbers when the calculation is done with paper and pencil; (3) the previously established mechanical associations among numbers known as addition, subtraction, multiplication and division; (4) The fact that the various articles are put together probably suggests the putting together of the numbers representing their value. That is, it acts as a stimulus to arouse the association process of addition rather than subtraction or multiplication or division; (5) the occurrence of the successive links of the mechanical associations of the numbers leading to the sum and the continued association of counting by units (pennies) until 100 is reached.

In brief the steps in the solution of a representative arithmetical problem would be as follows:

(1) Number concepts previously acquired.

(2) Ability to read, write or speak symbols for the numbers or objects to which they refer. The steps involved in these particular processes have been analyzed in the preceding chapters and hence need not be repeated here.

(3) Numerous connections previously established between the various numbers, which are known as the fundamental operations with numbers, fractions, and so on.

(4) A clue as to the particular connection to be made at successive points.

(5) The execution of the successive acts with the particular numbers involved.

(1) The number concept. This is partly acquired by the child before he enters school. It arises out of, and develops through, the endless experiences of the child in contact with objects and repeated occurrences of events. Some investigators of the genesis of the number idea believe it to arise by means of counting; others believe it to arise by means of grouping. The probability is that both factors contribute. What happens is substantially this: The child deals with various objects and learns through handling them that they are separate things and that there are several of them. As he continues to handle them, he established definite meanings and associations with each one. He furthermore develops the idea of number or quantity by finding possibly that one or more may sometimes be missing. He thus acquires a group perception and gradually widens it to the ability to count, which essentially consists of associating certain sounds and certain motor speech processes in speaking the sounds, with successive objects or acts, probably both, as the successive objects are touched and handled or at least looked at in turn. These elementary number ideas are then enlarged rapidly upon entering school by providing more extensive materials and by associating larger numbers with them.

(2) This step involves all the detailed and complicated elements enumerated in the reading and writing processes and as such need not be considered again as they contain no new elements. All the

other steps are specifically arithmetical processes and require detailed consideration.

(3) The establishment of associations among the numbers. The processes involved in the four fundamental operations of adding, subtracting, multiplying and dividing are pure association processes supplemented by experiences with illustrative material to show the meaning of the various operations. Ultimately and fundamentally they are established in the neural and mental machinery as associative links which, in the trained individual, become purely automatic. Six plus 2 equals 8, 6 minus 2 equals 4, 6 times 2 equals 12, and 6 divided by 2 equals 3, are pure association processes so that when the two numbers with a certain symbol between them appear and are spoken in succession, the appropriate last link is brought up thereby.

(4) The clue as to which process is to take place. The element in a given situation or problem that suggests which of the four operations at a given point shall take place is the clue which acts as a stimulus to arouse the appropriate associative reaction. In the illustration used, the fact of putting the articles together probably acts as the clue which starts the association of 10 plus 10 plus 45 plus 7 equals 72. Step (4) in this problem is what we ordinarily call reasoning. But fundamentally, arithmetical reasoning, as well as so-called reasoning in general, is essentially a matter of the perception of differences and similarities and a matter of selective association processes which are largely automatic in the perfected stages of arithmetical skill and largely trial and error in the early stages. The child associates the putting together of objects with addition; the taking away of objects from a group with subtraction; the putting together of groups of equal size with multiplication; and the taking away of groups of equal size with division. This at first probably develops through the multiplication of objects in these various ways. Something in any given situation or problem suggests one or the other of the four operations. Through indefinitely repeated situations, the correct operation is probably suggested as a matter of association. In a more complex situation in which the clue does not operate as automatically, the mind proceeds largely by trial and error by bringing up in turn different associations until the correct one arises, or until one arises that satisfies the circumstances.

In ordinary terminology, we call this process reasoning. It may seem as though we were reducing reasoning to mechanical association which naïvely we do not consider to be reasoning. Perhaps in a sense that is really what it amounts to. There probably is no such thing as reasoning in the sense of forcing one's thought processes into a given desired direction in a straight and direct line by sheer force of will. Reasoning, even that of the most original and inventive type, probably consists fundamentally in starting with a certain idea, desire, or problem, in short, with a stimulus, and in waiting for associations to arise and then in following out in turn by trial and error, one link after another, and in waiting for each one to bring up its links until a chain of successful links arises which satisfies the desire or which meets no opposition and which is then selected. Probably all that voluntary effort will do is to stimulate possibly a more rapid arousal of associative links, and to stir up by virtue of stimulating greater neural activity, such additional associations as ordinarily do not come up quite so easily. This is apparently what happens even in the inventive and most original type of thinking or reasoning. The statements of scientists and inventors bear witness to the error of the usual belief that original thinking goes straight to the goal with unerring step. Thus Jevons remarked:

"In all probability the errors of the great mind exceed in number those of the less vigorous one. Fertility of imagination and abundance of guesses at truth are among the first requisites of discovery."

And Faraday said:

"The world little knows how many of the thoughts and theories which have passed through the mind of a scientific investigator have been crushed in silence and secrecy by his own severe criticism and adverse examination; that in the most successful instances not a tenth of the suggestions, the hopes, the wishes, the preliminary conclusions have been realized." (Quoted from Lindley.)

Edison has described the invention of the electric lamp as follows:

"During all those years of experimentation and research, I never once made a discovery. All my work was deductive, and the results I achieved were those of invention, pure and simple. I would construct a theory and work on its lines until I found it was untenable. Then it would be discarded at once and another theory evolved. This was the only possible way for me to work out the problem. . . . I speak without exaggeration when I say that I have constructed 3,000 different theories in connection with the electric light, each one of them reasonable and apparently likely to be true. Yet only in two cases did my experiments prove the truth of my theory. My chief difficulty was in constructing the carbon filament. . . . Every quarter of the globe was ransacked by my agents, and all sorts of the queerest materials used, until finally the shred of bamboo, now utilized by us, was settled upon." (G. C. Lathrop, "Talks with Edison," *Harpers*, Vol. 80, p. 425.)

However, it must not be supposed that our analysis has more than sketched in its broadest outlines one of the most complex and difficult mental activities which human beings ordinarily perform. Indeed arithmetic is psychologically not a single activity but a group of fairly distinct activities, some of which are as unrelated to each other as arithmetic as a whole is to English. Thorndike found that the average arithmetic grades of the children in a certain school for a period of two and one-half years gave a correlation (Pearson) of .39 with the grades in English, and of .36 with those in geography. On the other hand, Stone tested 500 children on the four fundamental processes and arithmetical reasoning and obtained the following correlations:

Addition with subtraction	.50
Addition with multiplication	.65
Addition with division	.56
Subtraction with multiplication	.89
Subtraction with division	.95
Multiplication with division	.95
Arithmetical reasoning with addition	.28
Arithmetical reasoning with subtraction	.32
Arithmetical reasoning with multiplication	.34
Arithmetical reasoning with division	.36

These results are especially striking in their low correlation of the four fundamental processes with arithmetical reasoning. It is significant also that subtraction, which might be thought to be closely related to addition, turns out to be much more closely related to division, the correlations being .50 and .95 respectively. These results have added significance when it is remembered that all elements of subtraction or multiplication, for example, which are normally present in division fundamentals were carefully eliminated from the scores.

Our numerous previous observations of the specific nature of psychological activities, however, so far from making us surprised at Stone's results would rather lead us to expect still greater differentiation of the arithmetical processes. Howell found with regard to the combination $12-8=4$ and $12-4=8$ that knowing the one by no means assured knowing the other. Table 112 shows the total number of errors of 300 children in grades three to eight on eight pairs of complementary subtraction combinations:

TABLE 112

Total Errors		Total Errors	
12–4	14	12–7	17
12–8	0	12–5	0
13–5	8	15–8	17
13–8	15	15–7	0
11–6	10	11–9	15
11–5	0	11–2	0
11–7	22	17–9	12
11–4	5	17–8	1

Here there is evidently a strong tendency for the smaller of two amounts subtracted to yield a surprisingly greater degree of accuracy than its complement.

Now each of these numerous specialized arithmetical activities has a psychology more or less of its own. Most of these processes are as yet very imperfectly explored. The psychology of complementary processes just considered offers an excellent example. Other important ones are concerned with borrowing, carrying, and decimal points. Perhaps one of the most significant of all is a peculiar type of attention, or memory span, such as is involved in accurately keeping in mind and adding one to the tens at appropriate intervals while at the same time adding the units of a long column of digits; or, in subtracting, in which it is necessary to remember when borrowing has taken place so as to make the appropriate compensation. Among the specialized aspects of the psychology of arithmetic is the phenomenon of number preferences. The United States census reports show that when people are not certain of their ages they tend strongly to give even numbers and to avoid odd numbers. Phillips obtained introspections showing that people like the even numbers but feel uncomfortable at the thought of odd numbers. Jastrow in an unpublished investigation had university students estimate the number of steps of two short flights of stairs which all climbed nearly every day. The avoidance of the odd numbers and particularly of the prime numbers is very strikingly shown by Figure 75 where the results of both sets of judgments are combined. The avoidance of a given odd number is shown by comparing its frequency with the frequency of the numbers at each side of it. Eight is an especial favorite.

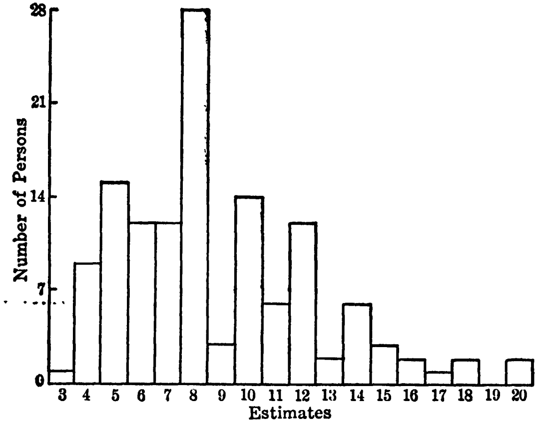

FIG. 75.—The number of persons whose estimates of the number of steps fall on various numbers showing a tendency to avoid odd and prime numbers.

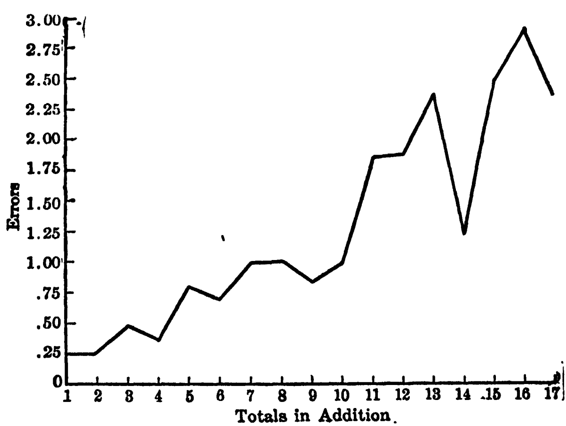

FIG. 76.—Showing the relative numbers of addition combinations amounting to the various totals. Adapted from Table 128.

These number preferences are chiefly significant to education in their genesis. Some indications as to their origin are incidentally furnished by two studies to be described presently, the results of which are represented graphically by. Figures 76 and 80. Figure 76 shows very clearly that there is a strong tendency for additions of digits which result in totals of odd numbers and especially of prime numbers, to be more difficult and thus to cause the learner more trouble than even-numbered totals. From quite a different source, namely, the difficulty of mastering number pictures (Figure 80), there comes evidence tending toward the same conclusion. Indeed it is possible that early experiences of a disagreeable nature involving these odd and prime numbers may have caused the unpleasant feelings reported by Phillips' subjects to be associated with them, and that this unpleasant feeling-tone may have inhibited to a certain extent the choice of odd. numbers in Jastrow's experiment according to the well-known tendency for the unpleasant to produce avoiding reactions. Moreover, the special difficulty encountered with prime numbers suggests that the difficulty of the various totals, apart from the factor of size, is closely related in its turn to indirect assistance derived from multiplication.

The Measurement of Efficiency in Arithmetical Operations

According to our analysis, the measurement of efficiency in arithmetical operations consists fundamentally in a measurement of the quickness and correctness with which the various operations are suggested and carried out under different conditions. Two types of measuring devices have been worked out. The first type consists of the well-known Courtis ('10) tests. Courtis prepared at the outset a series of eight tests now widely used, known as Series A, to measure the following operations: Addition, subtraction, multiplication, division, copying figures, reasoning without performance of the operations, fundamentals consisting of various combinations of the four fundamental operations, and reasoning, including the performance of the operations. The general principle of the tests consists in the selection of units of the material in each test of approximately the same type and difficulty throughout the test and of measuring efficiency by the number of operations attempted or done correctly in a given period of time. For example, the addition test is composed of a large number of combinations of single digits, such as 6 plus 3, 7 plus 1, etc. The measure of efficiency is the number of such additions made in one minute. Each of the other tests is constructed on the same general principle. The reasoning tests consist of separate problems of approximately equal length and difficulty, and the measure of efficiency is the number of problems attempted or solved correctly in a given number of minutes.

More recently Courtis prepared another set, Series B, which is confined to the four fundamental operations and contains consequently one test for addition, subtraction, multiplication and division. The chief difference between the tests in Series B and the corresponding ones in Series A is that the problems in the former contain larger numbers arranged in larger columns so as to introduce more complex elements.

The second general type of tests is constructed on the principle of a scale of problems of increasing difficulty. The author ('15) prepared a test for measuring ability in solving concrete problems, designated as Arithmetical Scale A. It is composed of a series of problems increasing in difficulty by determined steps of difficulty ranging from 0 to 15. Ability is measured not in terms of speed, but in terms of the highest point on the scale at which a pupil can solve problems correctly. The difficulty of the problem and the distances between the steps were determined by extensive experiments with pupils.

Stone ('08) prepared a set of problems whose difficulty was determined experimentally. The problems are, however, not arranged in scale form.

Woody ('16) prepared a series of tests on the scale plan for measuring ability chiefly in the fundamental operations. Each scale is composed of a series of problems arranged in the order of increasing difficulty.

Judd and Counts ('16) prepared for the survey of the Cleveland schools, fifteen sets of arithmetic tests, four for addition, two for subtraction, three for multiplication, four for division, and two for operations with fractions. The various sets for any one type of operation represented successively more and more complex stages of the operation.

Standards of attainment have been prepared by the authors of the various tests so that measurements of the abilities of pupils and schools can be compared directly with the respective norms for the various grades. These tests have been very useful in getting at some real facts concerning progress and attainment in arithmetic.

Figure 19, Chapter III, shows the range of individual abilities and the overlapping of abilities in the various grades in a certain school as measured by the author's arithmetical scale. The facts shown therein, although indicating enormous ranges of differences and overlappings, are substantially the same as those found in other subjects and hence are no longer surprising. Figure 20 exhibits the same facts for addition as measured by the Courtis tests.

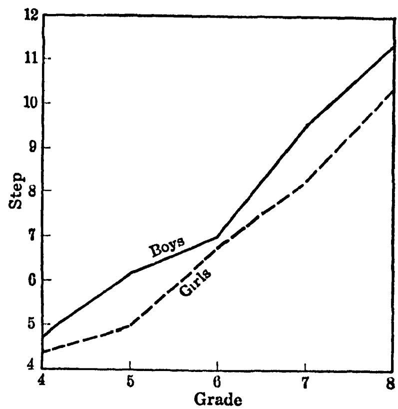

FIG. 77.—Difference between boys and girls in solving arithmetical problems as measured by the author s scale ('15).

Figure 77 shows the difference between the sexes and indicates that in arithmetical reasoning the boys surpass the girls. This is one of the few school abilities in which the boys are in the lead.

ECONOMICAL METHODS OF LEARNING ARITHMETICAL OPERATIONS

(1) **The Acquisition of the Number Concept.** The fundamental psychological materials which the child must acquire are ideas of numbers and quantities. He must learn what 1, 2, 3, 4, etc., mean. The practical question is, How may the pupil acquire these number concepts most economically? A considerable amount of experimental work has been done which bears either directly or indirectly on the processes by which the child develops definite notions of the significance of number and quantity. The investigations on the span of attention or apprehension as measured in terms of the number of objects that can be grasped simultaneously have chiefly

an indirect bearing. Such researches have been made by Cattell ('86), Messenger ('03), Burnett ('06), Dietze ('85), Warren ('97), Nanu ('04), Freeman, and others. The general result of these researches has been to show that the average person has a span of four, five or six objects or stimuli presented either simultaneously or in rapid succession to the sense of vision, hearing or touch.

Investigations bearing more directly upon the development of number concepts have been carried out by Lay ('98 and '07), Walsemann, Knilling, and others.

Pestalozzi was probably the first to attack the problem by devising for the purpose of school exercises his *Strichtabellen* (stroke-

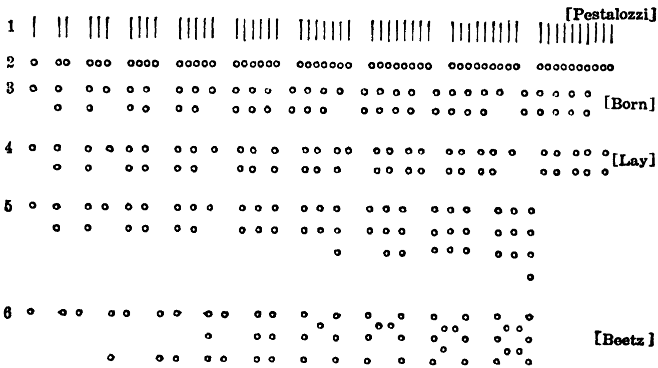

FIG. 78.—Number pictures or arrangements. After Howell (p. 154).

tables) which were the forerunners of more recent number pictures. These have been devised in two general forms: (a) those in which the dots are arranged in two parallel horizontal rows (prepared by Born, Busse, and Behme) and (b) those in which the dots are arranged vertically (prepared by Hentschel, Beetz, Sobelewsky, and Kaselitz). See Figure 78.

Extensive experiments have been made upon school children by Lay to determine the relative merits of the various devices and arrangements of the materials used for teaching apprehension of the number of objects. Lay investigated such problems as whether it is better for teaching numbers to present groups of stimuli or objects simultaneously or successively; whether it is better to present them in single rows or in double rows, in continuous rows or in quadrats (groups of four); how many objects children can apprehend at one time; what the effect is of distance between the objects; and so on.

The chief results have been summarized by Howell in the following Table 113, and graph, Figure 79.

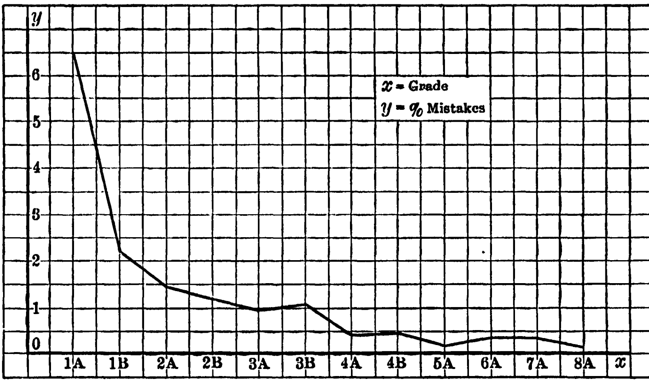

FIG. 79.—Curve of error for the number pictures 5 to 12 as a whole. After Howell ('14, p. 213).

TABLE 113. After Howell

Comparison of the Born, Beetz, and Lay pictures with one another and with the Russian Machine in the apprehension of the numbers 5 to 10. P = chances for mistakes. M = number of mistakes.

NUMBERS	LAY QUADRATIC PICTURES FIRST-YEAR PUPILS			BORN FIRST-YEAR PUPILS			WENDLING APP. (BORN) FIRST-YEAR PUPILS			BEETZ FIRST-YEAR PUPILS			RUSSIAN MACHINE FIRST-YEAR PUPILS		
	P	M	%	P	M	%	P	M	%	P	M	%	P	M	%
5	87	22	25	106	25	24	106	12	11	58	29	50	106	26	25
6	87	16	18	106	18	17	53	5	10	58	12	20	53	15	30
7	87	35	40	95	38	40	53	11	20	58	42	72	53	24	45
8	87	23	26	95	23	24	53	13	25	58	49	84	53	31	40
9	87	58	67	95	58	61	106	37	33	58	46	80	106	28	36
10	87	50	57	82	39	48	53	25	49	58	50	86	53	17	32
Total	522	204	40	579	201	35	424	103	24	348	228		424	141	
Total Training School	12,422	173	7	1,669	135	8	1,011	59	6						
Combined	2,944	377	13	2,248	336	15									

The outcome of these experiments is that the quadratic pictures are superior to other known pictures, particularly to the rows, and that the Lay and Born pictures are practically equal in merit.

Howell continued experimentation with quadratic number pictures similar to Lay's. The groups of dots used on the pictures ranged in size from three to twelve. Each card was exposed for 5 seconds. The investigation was carried out with pupils in an elementary school. Howell's results are shown in the curve of Figure 79. This curve shows a rapid drop in errors from the first grade to the second, and that pupils reach a high degree of certainty between the third and fourth grade in apprehending the numbers. Howell also found that certain numbers are not appre-

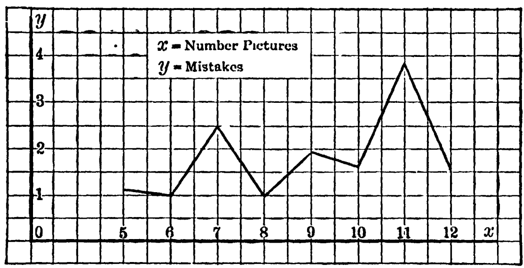

Fig. 80.—Number of mistakes in perceiving numbers by pupils in various grades. After Howell ('14, p. 217).

hended as accurately as others, the most difficult ones being seven and eleven, as shown in Figure 80.

(2) The Operations to be Learned. One of the first important problems in the economy of learning arithmetic, as in every school subject, is the determination of the operations that pupils should really master. Considerable attention has lately been given to this question with the result that the recent texts and courses of study have eliminated much material and have decidedly shifted the emphasis from certain topics to others.

Two methods of attack may be employed in determining the topics and the material to be learned in arithmetic. According to the one method, we might obtain the consensus of many expert opinions as to the topics that should be included and the relative amount of emphasis on each. According to the other method, we might proceed to gather a large number of the arithmetical problems and operations actually involved in the occupations and professions of all classes of people, and to determine on the basis of such a collection, the sort of material that should be taught. Some results have been obtained by both methods.

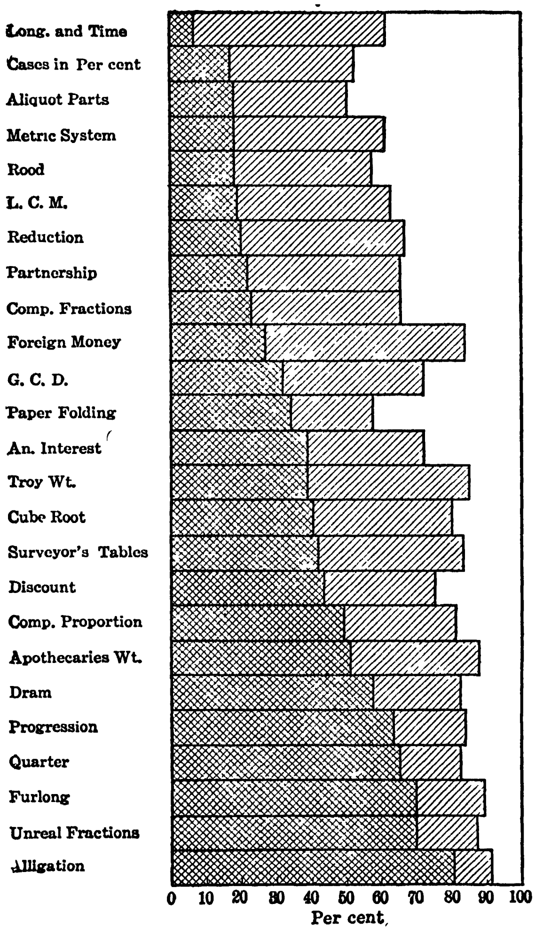

FIG. 81.—Percentage of superintendents in 830 cities who favor elimination of the various topics represented by checked surface and those who favor "less attention" are represented by shaded surface. After Jessup and Coffman ('15).

Jessup ('15) and Coffman obtained an expression of opinion from superintendents of 830 cities with a population of 4,000 or over as to which topics should be eliminated, which should receive less emphasis, and which should receive more emphasis than they do at present. Their results are shown in Figure 81 and Figure 82.

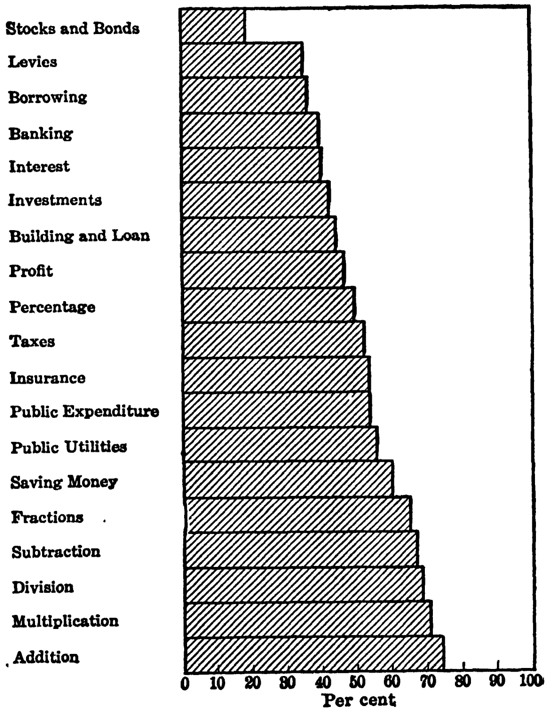

FIG. 82.—Percentage of superintendents in 830 cities who favor giving more attention to each of these topics. After Jessup and Coffman ('15).

As a result of these returns Jessup and Coffman recommend the elimination of the following topics from the elementary course of study:

"Apothecaries' weight, alligation, aliquot parts, annual interest, cube root, cases in percentage, compound and complex fractions of more than two digits, compound proportion, dram, foreign money, folding paper, the long method of greatest common divisor, longitude and time, least common multiple, metric system, progression, quarter in avoirdupois table, reduction of more than two steps, troy weight, true discount, unreal fractions."

Furthermore, they recommend greater attention to such topics as:

"Time saved through the omission of the material mentioned in the foregoing may be wisely devoted to the study of social, economic and arithmetical issues involved in such facts as saving and loaning money, taxation, public expenditure, banking, borrowing, building and loan associations, investments, bonds and stocks, tax levies, insurance, profits, public utilities, and the like."

G. M. Wilson ('17) collected 5,036 problems from 1,457 persons, representing practically all varieties of occupations and professions. He then classified these problems according to the type of operation involved, and the number of problems of each type as shown in Table 114.

TABLE 114. After Wilson ('17)

Addition	1-Pl	30	Accounts	251
	2-Pl	706	Addition of Fractions	3
	3-Pl	748	Amount	11
	4-Pl	193	Area	1
Over	4-Pl	65	Average Weight	14
Total		1742	Banking	18
			Board Measure	12
			Cancellation	26
Multiplication	1-Pl	1660	Capacity	10
	2-Pl	904	Circular Measure	1
	3-Pl	195	Cubic Measure	56
	4-Pl	17	Debts	56
Over	4-Pl	3	Decimals	4
Total		2779	Discount	5
			Division of Fractions	1
Subtraction	1-Pl	40	Dry Measure	5
	2-Pl	407	Exchange	5
	3-Pl	406	Insurance	10
	4-Pl	167	Interest	66
Over	4-Pl	65	Liquid Measure	14
Total		1085		

TABLE 114—Continued

Division	1–Pl	334
	2–Pl	319
	3–Pl	121
	4–Pl	48
Over	4–Pl	5
Total		839
Fractions	1/(2–5)	534
	1/(6–10)	12
	1/(10 plus)	4
	(1 plus)/(1–5)	86
	(1 plus)/(5 plus)	60
Total		1035
United States Money	1–Pl	23
	2–Pl	2982
	3–Pl	1714
	4–Pl	550
Over	4–Pl	247
Total		5516

Making Change	3
Measuring	21
Percentage	217
Plastering	2
Practical Measurement	79
Profit and Loss	16
Proportion	5
Receipts	1
Square Measure	27
Taxes	6
Time Measure	13
Buying	3128
Selling	646

"The problems solved in actual usage are brief and simple. They chiefly require the more fundamental and more easily mastered processes.

"In actual usage, few problems of an abstract nature are encountered. The problems are concrete and relate to business situations. They require simple reasoning and a decision as to the processes to be employed.

"The study justifies careful consideration of the following question: After the development of reasonable speed and accuracy in the fundamentals and the mastery of the simple and more useful arithmetical processes, should the arithmetic work not be centered largely around those problems which furnish the basis for much business information?" (Wilson.)

W. S. Monroe ('17) has compiled the problems in four text-books and classified them according to the types of operation involved and then compared the frequency of these types of problems with the number of workers in the different occupations. His preliminary report states:

"In the first place, out of a total of 1,023 types of practical problems found in four text-books, 720, or 71%, occur in occupational activities.

"A study of the frequency with which type problems occur reveals a significant fact; viz., the frequency ranges from one to 434.

"This wide variation in frequency shows that the authors of our text-books are far from being in agreement on the type problems of arithmetic. Only one author out of four has recognized 511 out of 1,023 type problems and 140 type problems have received the recognition of only two authors out of four." (Monroe.)

(3) **Length of the Class Period.** J. M. Rice made an investigation of efficiency in arithmetic after the general plan of his investigation of spelling. He tested some 6,000 pupils in eighteen different schools in seven cities. His results are exhibited in Table 115, which

"Gives two averages for each grade as well as for each school as a whole. Thus, the school at the top shows averages 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."

TABLE 115. After Rice ('02)

Averages for individual schools in arithmetic

	School	4th Year Result	4th Year Principle	5th Year Result	5th Year Principle	6th Year Result	6th Year Principle
City III	1	68.4	76.7	79.5	82.5	79.3	80.3
City I	1	72.7	81.4	84.7	88.8	80.4	81.5
City I	2	....	....	80.3	87.1	80.9	83.4
City I	3	54.5	66.0	74.7	78.3	72.2	74.0
City I	4	60.0	73.3	70.8	79.3	69.6	72.2
City II	1	81.3	87.7	78.2	85.6	71.2	75.3
City III	2	70.1	79.9	53.6	60.0	43.7	45.0
City IV	1	70.5	76.4	73.2	77.7	58.9	60 4
City IV	2	62.9	72.8	70.5	76.8	59.8	63.1
City IV	3	59.8	72.7	65.3	73.5	54.9	58.1
City IV	4	53.5	62.2	53.5	65.7	42.3	45.1
City V	1	38.5	46.3	67.0	71.0	44.1	48.7
City VI	1	28.1	31.7	38.1	44.2	68.3	71.3
City VI	2	41.6	52.9	45.3	52.1	46.1	49.5
City VI	3	36.8	54.2	55.0	62.8	34.5	36.4
City VII	1	59.3	69.3	53.7	63.1	35.2	37.7
City VII	2	47.4	55.2	65.4	71.4	35.2	38 7
City VII	3	41.1	58.0	37.5	44.5	27.6	33 7
General average		59.5	69.9	69.4	75.5	60.7	63.2
Per cent of mechanical errors			14.8		8.1		3.9
Number of pupils examined			1,422		1,593		1,285

TABLE 115. After Rice ('02)—Continued

Averages for individual schools in arithmetic—Continued

	School	7th Year Result	7th Year Principle	8th Year Result	8th Year Principle	School Average Result	School Average Principle	% of Mechanical Errors	Min. Daily	When Taken
City III	1	81.1	82.3	91.7	93.9	80.0	83.1	3.7	53	A. M.
City I	1	64.2	67.2	80.9	82.8	76.6	80.3	4.6	60	A. M.
City I	2	43.5	50.9	72.7	79.1	69.3	75.1	7.7	25	A. M.
City I	3	63.5	66.2	74.5	76.6	67.8	72.2	6.1	45	A. M.
City I	4	54.6	57.8	66.5	69.1	64.3	70.3	8.5	45	P. M.
City II	1	33.6	35.7	36.8	40.0	60.2	64.8	7.1	60	P. M.
City III	2	53.9	56.7	51.1	53.1	54.5	58.9	7.4	60	P. M.
City IV	1	31.2	34.1	41.6	43.5	55.1	58.4	5.6	60	A. M.
City IV	2	....	....	22.5	22.5	53.9	58.8	8 3	..	P. M.
City IV	3	35.2	38.6	43.5	45.0	51.5	57.6	10.5	60	A. M.
City IV	4	16.1	19.2	48.7	48.7	42.8	48.2	11.2	..	P. M.
City V	1	29.2	32.5	51.1	58.3	45.9	51.3	10.5	40	P. M.
City VI	1	33.5	36.6	26.9	30.7	39.0	42.9	9.0	33	A. M.
City VI	2	19.5	24.2	30.2	40.6	36.5	43.6	16.2	30	P. M.
City VI	3	30.5	35.1	23.3	24.1	36.0	42.5	15.2	48	P. M.
City VII	1	29.1	32.5	25.1	27.3	40.5	45.9	11.7	42	A. M.
City VII	2	15.0	16.4	19.6	21.2	36.5	40.6	10.1	75	P. M.
City VII	3	8.9	10.1	11.3	11.3	25.3	31.5	19.6	45+	P. M.
General average		39.4	42.5	49.4	51.9	55.7	60.6	8.1		
Per cent of mechanical errors			7.3		4.8					
Number of pupils examined			974		689	Total, 5,963				

With reference to the factor of time in relation to efficiency in arithmetic, Rice concludes thus:

"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%—was practically the same as it was in the one where the highest average—80%—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%. 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% and 67%, respectively, the school that gave only twenty-five minutes succeeded in obtaining an average of 69%. 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."

Stone ('08) made a similar investigation, testing some 6,000 pupils in the 6th grade in twenty-six school systems. He reports results practically identical with those of Rice, namely, that while the amount of time devoted to arithmetic in different schools varied from 7% to 22% of the total school time, yet a comparison of time expenditure with the efficiency attained showed, according to his interpretation, that time plays a negligible part.

These results and inferences are interesting and valuable but they cannot be interpreted with absolute assurance. The various factors coöperating or counteracting are so intricate that a more careful isolation of the effect of the time element is necessary. In general, the same criticism made in connection with Rice's and Cornman's investigations of spelling applies here. Dependable conclusions could be reached only by an experimental procedure similar to the one there suggested.

The findings of Rice and Stone probably represent correctly the situations in the schools examined. A possible explanation of the fact that the schools giving more time to arithmetic did not obtain on the whole any higher efficiency than those devoting less time to it may, perhaps, be sought in the likelihood that the schools giving longer periods of time may not have worked as intensively and used their time to as good advantage as the schools devoting less time.

(4) **The Effect of Various Environmental Factors.** Both Rice and Stone massed their results with reference to ascertaining the effect of such factors as the home environment of the pupils, size of classes, age of pupils, the time of day of the test, amount of home-work required of the pupils, method of teaching, teaching ability, the course of study, the superintendent's training of the teachers, etc. Rice reports that none of the factors had any influential part in producing efficiency in arithmetic. The results are open to the same criticism of complication of factors as were pointed out previously. It seems quite improbable that these elements played no part. It is rather a question of more rigorous isolation of the effect of different factors. Stone, for example, found that the correlation of excellence in the course of study, as rated by judges, with efficiency in arithmetical reasoning was .56, and with efficiency in fundamentals .13.

That environmental factors, and perhaps particularly method and spirit of teaching, do make important differences in the attainments of pupils is shown clearly in such results as those exhibited in Table 116 which gives the distribution of class averages of the grades in sixteen different schools as measured by the author's Arithmetical Scale A.

TABLE 116

Average scores attained in various schools as measured by Arithmetical Scale A (Starch)

Grades	3	4	5	6	7	8
City A						9.7
B School 1						13.1
2			7.2	10.4	10.6	11.2
C School 1	5.1	5.9	7.2	9.2		
2	3.9	5.6	6.9			
3	3.9	5.3	5.6	7.5	9.2	12.6
G School 1			9.0	10.9	11.6	14.5
2			8.9	12.0	13.0	13 7
3		7.5	10.2	9.2	10.9	11 5
4			10.0	10.6	11.3	
I School 1	5.1	6.0				
2			6.2			
3				10.0	10.2	11 0
L School 1	4.6	5.8	8.5	9.8	11.9	14 0
2		4 6	7.4			
3	6.9	8.5	11.3			
4		6.6	8.8			

Thus we note that the best eighth grade attained an average of 14.5 as compared with the poorest one which attained an average of only 9.7. Such differences would not be surprising if they were the scores of individual pupils. They are, however, the mean scores of whole classes. It is quite unlikely that the hereditary differences of the groups as wholes differ so much from one another. It seems quite probable that the environmental circumstances, and chief among them the teacher and the attitude of the learner, were mainly responsible for the ultimate differences in achievement.

Similar results have been reported by Judd for the fifth and eighth grades in ninety schools in Cleveland, as measured by his Test A in simple addition, Figure 83. The best fifth grade made an average score nearly three times as high as the poorest fifth grade, and the best eighth grade made a score nearly twice as high as the poorest eighth grade.

(5) **Drill in Fundamental Operations.** Various methods of drill in the fundamental operations have been devised.

Studebaker, Assistant Superintendent of Schools at Des Moines, has prepared a series of drill cards. The various combinations of numbers in fundamental operations are given on one side of the card. Below each example there is an opening through the card in which the pupil may write his answer on the sheet of paper placed underneath the card. The pupil works as rapidly as he can within a certain limit of time. Then he turns the card over and places it again over the sheet of paper containing his answers so

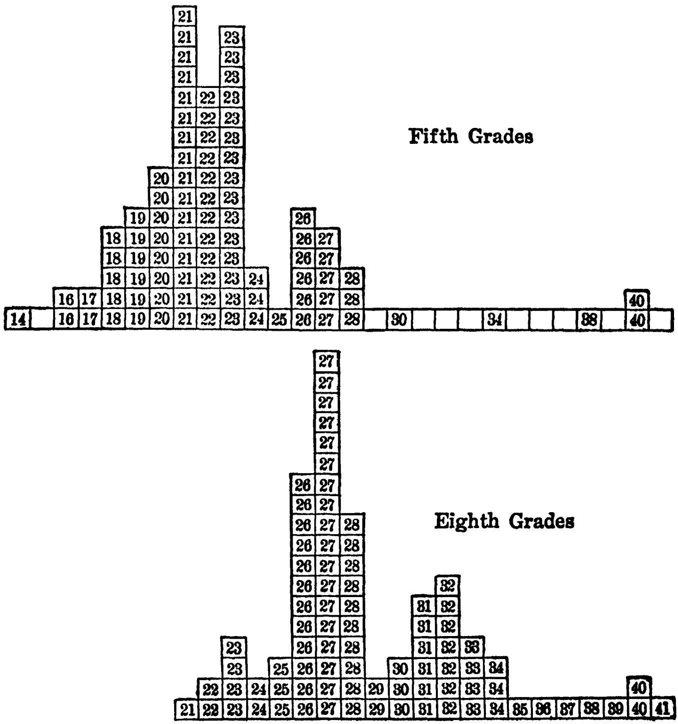

FIG. 83.—Median scores of the 5th grades and of the 8th grades in 90 schools in simple addition. After Judd ('16, p. 112).

that they can be seen in the openings of the card and compared with the correct answers printed on that side of the card.

This plan of drill work has a number of advantages, such as an incentive to rapid and accurate work, immediate self-checking of the answers, and so on.

A considerable number of careful experimental studies on the influence of drill are now available and without exception they show drill to be distinctly valuable. Thorndike had nineteen university students practice adding 48 10-digit columns of figures daily for seven days. While the work required on the average less than an hour in all, there was an improvement of 29% over the original rate.

J. C. Brown performed two elaborate comparative experiments to determine whether children under controlled school conditions profit more by giving a small part of each class period to drill or by spending the entire period in ordinary routine work in arithmetic. In each experiment the children were first tested with the Stone Arithmetic tests and then divided into two groups of equal ability on the basis of their performance in the tests. One group was given the special drill as a part of the regular class work while the other did the class work as usual. At the conclusion of the drill, both were tested again by the Stone tests to see which had made the greater gain. In the first experiment 51 children from the sixth, seventh, and eighth grades were used. They averaged thirteen and one-half years of age. Drill on the four fundamental operations was given to one-half of the group for the first five minutes of each class period of twenty-five minutes. About half the drill was oral and half was written. The drill lasted thirty periods. In the second experiment 222 children were used and the drill was given for twenty periods. The results of the two experiments are given in parallel columns in Table 117. In each case section I received the drill and section II received the regular class work. The pupils did not know that any experiment was in progress.

TABLE 117

Section	Per Cent of Improvement of Second Test over First in	Per Cent First Experiment (51 Children)	Per Cent Second Experiment (222 Children)
I.........	Number of problems worked	21.2	16.9
II........	" " " "	9.8	6.4
I.........	Fundamentals, Addition	33.4	18.5
II........	" "	11.8	6.8
I.........	Fundamentals, Subtraction	36.9	32.0
II........	" "	13.1	11.9
I.........	Fundamentals, Multiplication	30.0	24.1
II........	" "	13.7	10.9
I.........	Fundamentals, Division.......	28.0	34.2
II........	" "	19.3	15.4
I.........	Total number of points made	32.0	24.2
II........	" " " " "	14.7	9.4
I.........	Number of points made on the first six problems (averaged)	5.8	11.7
II........	Number of points made on the first six problems (averaged)	2.4	−1.8

In both the experiments there was a decided advantage in using a part of the recitation period for drill. In the first experiment, the drilled group gained about twice as much as the undrilled group; while in the second experiment the drilled group improved about two and one-half times as much as the undrilled group. The sixth grade gained the most (35%) and the eighth the least (13.8%). In order to determine whether group I had gained on fundamentals at the expense of reasoning, both groups were tested in arithmetical reasoning before and after the drill. Here again the drilled group did better, making a gain of 6.3%, while the undrilled group gained only 3.0%. This last factor is interesting in the light of the small amount of connection between fundamentals and arithmetical reasoning pointed out above as well as the small amount of transfer of one arithmetical process to another (Chapter XIV). Since the improvement in reasoning, which had not been drilled at all was almost exactly the same proportionately as the processes which were drilled, it suggests that the drill had a tonic effect upon the remainder of each recitation period following the drill, to which much of its value was due.

In order to discover the permanency of the effects of drill, Brown tested both groups once more after a twelve-weeks vacation and found that the drilled group was also superior in retention, having lost .2% while the undrilled group had lost 2.29%.

The same experiment was repeated by F. M. Phillips. He had 69 children for subjects and gave to one group drill in fundamentals and in reasoning, both oral and written, for eight weeks. Neither teachers nor students knew the purpose of the tests. He found that, "The improvement in fundamentals of the combined drill groups was 15% greater than that of the non-drill groups. In reasoning, the drill groups improved 50% more than the non-drill groups. . . . The greatest gains were made in the sixth grade and the least in the eighth." Almost all the gain on fundamentals was in multiplication.

Mary A. Kerr under the direction of Haggerty reported an experiment carried on for six weeks at Bloomington, Indiana, on the effects of five minutes of drill in addition at the beginning of each class period. The drill was begun by adding five three-place numbers per column, which were gradually increased to nine three-place numbers per column. Four hundred and twenty-three children took the drill. Table 118 shows the average performance on the Courtis tests, Series B, before the drill began and at its conclusion in June and, for comparison, the May scores of the best twenty Indiana cities for the previous year.

TABLE 118

Grade	Attempts Bloomington		Highest Median Scores of 20 Ind. Cities (May)	Rights Bloomington		Highest Median Scores of 20 Ind. Cities (May)	Per Cent Accuracy Bloomington		Highest Median Scores of 20 Ind. Cities (May)
	Feb.	June		Feb.	June		Feb.	June	
6B ...	8.7	9.7	8.9	5.5	6.9	5.6	64	72	65
6A ...	9.0	10.5		5.3	7.5		59	71	
7B ...	9.7	10.8	9.4	5.6	8.1	6.4	60	75	68
7A ...	9.8	11.8		6.0	8.4		62	71	
8B ...	11.4	12.0	10.3	6.9	9.3	7.2	61	78	69
8A ...	11.5	13.7		6.3	10.4		55	76	

The decided advance in each grade and the great superiority to the best twenty Indiana cities in each test bears eloquent testimony to the value of drill in addition fundamentals.

Supt. Herman Wimmer of Rochelle, Illinois, conducted a series of comparative experiments on the effects of drill in arithmetic under various conditions. Each experiment lasted six weeks. The time spent in drill was in all cases subtracted from the regular class time in arithmetic. The Courtis tests, Series A, were given before and after each experiment. In Experiment I, the drill group was rather miscellaneous, grade 5 being drilled five minutes per day (probably in fundamentals, though it is impossible to tell from Wimmer's account), grade 6E five minutes per day, three-fifths on reasoning and two-fifths on fundamentals, grade 6W fifteen minutes in a single period per week on reasoning and fundamentals in the same ratio as 6E. In Experiment II, two sections of the sixth grade, which had equal ability as shown by tests, were drilled five minutes daily, one for speed and the other for accuracy. In Experiment III, the seventh grade was drilled five minutes daily on reasoning, while the eighth grade was drilled five minutes daily on fundamentals. The results of all three experiments are shown by percentages of gain in Table 119. As in the previous experiments, drill as such is shown to have a decided value though much more for reasoning than for fundamentals, the advantage of the drilled groups being 32.7% and 4.4%, respectively. In Experiment II, drill for speed is seen to have a distinct advantage over drill for accuracy. Here the gain is considerably more in the fundamentals than in reasoning. In Experiment III, we find that the class drilled on reasoning gained very largely in reasoning alone, while the other class trained in fundamentals gained almost exclusively in fundamentals. This presents a contrast to Brown's experiment where, it will be remembered, drill in fundamentals showed as much gain in reasoning as in fundamentals themselves.

TABLE 119

Wimmer's results

Problem	Experiment I Drill vs. No Drill			Experiment II Drill for Speed vs. Drill for Accuracy			Experiment III Drill in Reasoning vs. Drill in Fundamentals	
Grades Used as Subjects / Number of Subjects / Type of Activity	5th 6th / 79 / Drill	7th 8th / 70 / No Drill	Difference in Favor of Drill	6E / 22 / Drill Speed	6W / 22 / Drill Accuracy	Difference in Favor of Drill for Speed	7th / 35 / Drill Reasoning	8th / 35 / Drill Fundamentals
All tests, attempts and rights	33.3	16.2	17.1	11.1	8.8	2.3	9.1	8.6
Fundamentals	15.6	11.3	4.4	12.9	6.9	6.0	4.7	13.3
Reasoning, attempts and rights	58.0	25.1	32.9				18.6	.4
Reasoning, rights only	68.1	35.4	32.7	11.0	0.6	1.3	19.1	.8

(6) The Optimum Distribution of Drill. A number of experiments have been conducted to determine the most economical distribution of time, that is, the most economical duration of drill periods in arithmetic. T. J. Kirby carried out such experiments on a large scale for both addition and division. He used special blanks for the practice in both experiments. In addition fundamentals there was a beginning and a final test each of fifteen minutes. Between these periods forty-five minutes of drill were variously distributed. The following table shows the distribution of the periods, the median initial ability of each group in examples correct, and the gross gains as measured by three different methods of calculation. Seven hundred and thirty-two fourth-grade children were used as subjects.

TABLE 120. After Kirby

Group	No. Subjects	Initial Test Period	Distribution of Intervening 45 Minutes	Final Test Period	Median Initial Ability (Examples Correct)	Gain Due to Drill: Av. Gross Gain of Individuals	Med. Gross Gain of Individuals	Av. of Med. Gross Gains
I	194	15 min.	2 periods 15 min.	22½ min.	22.9	11.0	9.5	10.2
II	104	15 "	3 " 15 "	15 "	25.4	13.6	11.0	9.6
III	205	15 "	7 " 6 " and one 3 "	15 "	21.0	10.7	9.6	9.4
IV	229	15 "	21 periods 2 " and one 3 "	15 "	25.1	16.1	12.6	13.9

When accurate correction had been made for differences in initial ability, the gains were then in proportion 100, 121, 101, and 146½ respectively. There was no very distinct tendency observable here except that the short periods of two minutes yielded a distinct advantage over the rest. Unfortunately for the consistency of these results the next shortest period (6 min.) yielded nearly the least gain of all.

The same general plan was followed by Kirby in the experiment on division. Six hundred and six children from the second half of grade three and the first half of grade four were used. The following table shows in detail the various distributions of the drill together with the results by each method.

TABLE 121. After Kirby

Group	No. Subjects	Initial Test Period	Distribution of Intervening 40 Min. of Drill	Final Test Period	Median Initial Ability (Examples) Correct	Gain Due to Drill: Av. Gross Gain of Individuals	Gain Due to Drill: Med. Gross Gain of Individuals	Gain Due to Drill: Av. of Median Gross Gains of Classes
I ...	204	10 min.	2 20 min. periods	10 min.	38 4	25.1	22.6	20.6
II...	209	10 "	4 10 " "	10 "	33 4	25.5	23.5	25.1
III..	193	10 "	20 2 " "	10 "	41 4	42.6	40 4	44 7

When inequalities of initial ability had been removed, the gains were found to be in the proportion of 100, 110½, and 177. Thus we find here very consistent and decided advantage in favor of the shorter drill periods. Unfortunately in each of these experiments it is impossible to tell how much of the gain in the shorter practice periods was due to spontaneous practice outside the class. It is needless to say, that the children were not permitted to take any of the practice cards away from the class.

Kirby's experiments as a whole both in addition and division showed great improvement. Addition with a practice period of 60 minutes yielded an improvement of 48%, while division with a practice period of 50 minutes yielded an improvement of 75%. Accuracy was not disturbed in addition, but in division it improved 2.6%.

Kirby also investigated the permanence of the improvement resulting from the drill. He found that from June to September fourth-grade children lost 17% of the ability possessed in June and required 58% as much time to regain the efficiency which they possessed the preceding year. In division there was a loss of 21%.

It required 60% as much time to recover the preceding year's efficiency.

Hahn and Thorndike repeated the addition part of Kirby's experiment. Each grade was divided into two sections, section B receiving periods exactly half as long as section A. All received a total of 90 minutes of drill, which was preceded and followed by a 15-minute test as in Kirby's investigation. Table 122 shows the distribution of drill for the various groups, the initial ability, the gain, and the advantage of group A according to two different methods of scoring the results.

TABLE 122. After Hahn and Thorndike

Grade	Number of Subjects	Length of Practice Period	Score Based on Rights only			Score Based upon the Rights + One-Half the Wrong Answers		
			Average Initial Score	Average Gain	Advantage in Favor of Group A	Average Initial Score	Average Gain	Advantage in Favor of Group A
7 A ...	10	22½ min.	25.9	25.7	8.2	32.2	23.8	6.2
7 B ...	16	14¼ "	26.0	17.5		33.1	17.6	
6 A ...	13	20 "	16.3	10.7	.0	22.0	11.3	−3.4
6 B ..	12	10 "	17.4	10.7		22.2	14.7	
5 A ..	9	15 "	13.5	11.4	1.4	17.2	15.3	−2.3
5 B	12	7½ "	14.8	10.0		19.5	17.6	

There is no clearly defined advantage for either the long or the short period as was the case with Kirby's experiment on drill in adding. What little tendency there is, however, is in favor of the longer periods.

Superintendent Wimmer, in connection with his drill experiments previously reported, also investigated the problem of economy of long and short periods of practice. The drill, three-fifths on reasoning and two-fifths on fundamentals, was given to one group for five minutes at the beginning of each class period. The other group received one 15-minute period per week. Drill lasted for six weeks. The results are shown in the following table. They are distinctly in favor of the longer drill period despite the fact that only three-fifths as much time was spent by this method.

TABLE 123

After Wimmer

Problem	Favorable Distribution of Time for Drill		
Grades Used Number of Subjects Type of Activity	6E 22 5 Min. Drill Daily	6W 22 15 Min. Drill Weekly	Difference in Favor of Drill once per Week
All tests, attempts and rights...	34.3	45.1	10.8
Fundamentals................	14.7	18.1	3.4
Reasoning, attempts and rights..	67.2	83.5	16.3
Reasoning, rights only.........	75.9	102.1	26.2

This is particularly the case with reasoning. The rather slight advantage of the long periods on fundamentals in connection with the uncertain indications of Hahn and Thorndike's results and the opposite finding of Kirby suggests that there is little or no advantage in the distribution of time in arithmetical drill on fundamentals but that the longer periods are more favorable for drill in reasoning.

(7) **Special or Economical Methods of Drill.** A number of special methods of giving drill in arithmetic have been advocated and used. That of Studebaker has already been noticed. Courtis has also published practice pads for drill purposes. Flora Wilbur undertook to determine experimentally the value of this kind of drill at the Fort Wayne, Indiana, training school. Two classes of 14 children each were divided each into two groups of equal ability on the basis of the Courtis tests, Series B. One section of the fifth grade received four and one-half minutes of drill with the pads at the beginning of each class period, and one section of the sixth grade received similar drill for four minutes. The remaining sections received the regular class work. The experiment lasted from September to May. The results are shown in percentages of gain in the following table. The drill was clearly of value in both grades and in all four processes.

TABLE 124. After Wilbur

	Grade Five						Grade Six					
	Speed			Accuracy			Speed			Accuracy		
	No Practice	Practice	Gain of Practice over no Practice	No Practice	Practice	Gain of Practice over no Practice	No Practice	Practice	Gain of Practice over no Practice	No Practice	Practice	Gain of Practice over no Practice
Addition .	30	43	13	30	37	7	43	60	17	12	20	8
Subtraction . . .	27	54	27	29	49	20	28	49	21	2	21	19
Multiplication . .	12	148	136	26	31	5	51	88	37	18	20	2
Division . .	0	144	144	43	42	−2	76	129	53	28	34	6

Division and multiplication profited most and addition least, as is usual in such experiments. The real question remains, however: Is drill with the Courtis practice pad more or less efficient than drill as ordinarily given?

Kirkpatrick performed two comparative experiments to determine the relative economy of various methods of memorizing multiplication tables. His subjects were twenty normal school men, divided into groups of equal ability. As these subjects knew the ordinary tables, he had them learn the products of 7 multiplied by all the prime numbers between 17 and 53. One group simply memorized the table by rote for the first five days, then spent the periods of the next five days in writing down the answers on a blank with a card containing the table before them for reference. The other section spent the periods of all ten days working on the blank with the table for reference. The time consumed during the five days of memorizing was about an hour. A test at the end of the experiment showed that, in a period of two minutes, the practiced group put down 46.2 answers while the memorizing group put down 40.9, thus showing a distinct advantage for the practiced group.

A second experiment was performed with two groups of normal school students of equal ability, twenty-five in each group. One group practiced with the keys and blanks as in the first experiment while the other group spent the same amount of time multiplying out the products as needed. The experiment extended over eight days. When tested at the end of the experiment, those using the key put down 25.4 answers in two minutes, while the computers wrote down 44.3 examples, showing a decided advantage for the computation method. He concludes that tables should be learned by use rather than by memorizing.

Conrad and Arps ('16) investigated the effect of suppressing articulatory movements upon the effect of drill in rapid adding. They divided sixty-four high school students into equal groups of equal ability. The students were then given eight periods of drill in rapid addition of columns. The pupils in one group were permitted to add in their ordinary way which involved a great deal of articulation or inner speech. The other group was cautioned repeatedly and emphatically to "think results only." The former was called the traditional method and the latter the economical method. The percentages of gain by the two methods were as follows:

The traditional method gained in attempts 8.5% and in rights 2.5%. The economical method gained in attempts 34.4% and in rights 30.9%. This gave an advantage in favor of the economical method of 25% in attempts and of 33.4% in rights. These results came out almost startlingly in favor of "thinking results only." The evil effects of articulation and lip movements have been noticed in connection with reading (page 287). It is probable that the cause is the same in both cases.

P. B. Ballard investigated the comparative efficiency of the "equal addition" method and the "decomposition" method in subtraction.

"In the equal addition method the compensation is made—accounts are squared—at the very first number dealt with after the minuend has been disturbed. In subtracting 37 from 85, after taking 7 from 15 the disturbed relationship of difference between minuend and subtrahend is immediately restored by increasing the 3 tens to 4 tens. In the method of decomposition, however, it is the 8, the second figure dealt with, that has to be changed to restore the balance. If the minuend figure is zero, the balancing of accounts is still longer deferred."

Ballard gave tests in fundamentals to 71 English schools of which 23 had been taught subtraction by the method of equal addition while the rest had been taught it by the method of decomposition. While there was little difference in the average ability of the two groups in the other three fundamental operations, there was a very striking superiority in the score for subtraction of the equal addition group. At 13 years of age it amounted to over 10% and at earlier ages it amounted to over 40%. (See Figure 84.) Inasmuch as the decomposition, or less efficient method, is the one in general use in this country, it is evident that this matter deserves careful attention.

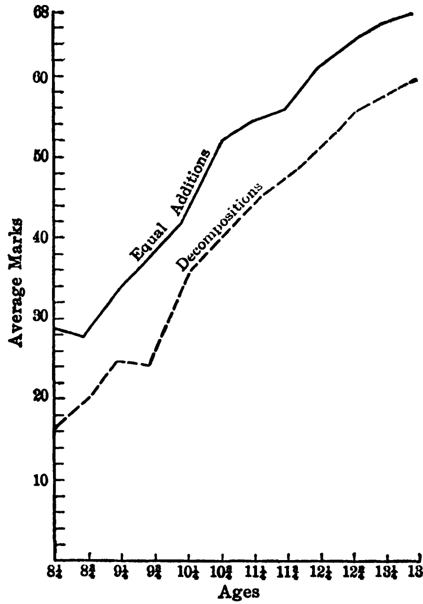

FIG. 84.—Showing superiority of teaching subtraction by the "equal additions" method. After Ballard ('15).

Mead and Sears performed two experiments in comparative economy of methods in arithmetic. The first was to compare the efficiency of the ordinary "take away" method of subtraction which involves the learning of an entire subtraction table, as compared with "addition" subtraction which permits the use of the addition table, thus saving the learning of an entire table. In the first the formula is "8 minus 2 equals what?", in the second, "2 plus what equals 8?" Two second-grade classes of approximately equal median ability as indicated by the Courtis tests were taught subtraction by the respective methods thirty minutes per day for four months, all other factors being equalized as fully as possible. Tests given periodically throughout the experiment showed that at the outset the addition method was superior but the "take away" method gradually overtook it until at the end of the four months the "take away" method was superior by 4.5 points which was nearly one-third of the final median score made by the addition group. This difference disappeared, however, when both groups were tested on longer examples.

The second experiment was to compare the multiplicative method of division. The formula of the first is, "Five into twenty how many times?", that of the latter, "Five times what equals twenty?" Two third-grade classes were used as subjects in this experiment. Other conditions were similar to the experiment on addition. In this experiment, the final test on combinations revealed the fact that the multiplicative class stood 4.3 points above the "into" class, which was about one-fifth of the final score of the "into" class. This difference disappeared, however, when the class was tested on longer examples just as that in addition noted above.

J. A. Drushel investigated the relative efficiency of two methods of determining the position of the decimal point of the quotient in the division of decimals. The rule of method A, the older one, is: "There are as many places in the quotient as those in the dividend exceed those in the divisor." The rule of method B, the newer of the two, is: "First render the divisor an integer by multiplying both dividend and divisor by 10 or some power of ten. Then proceed as with integral divisors." A short test in division of decimals was given to 576 freshmen at Harris Teachers' College. Of these, 507 had studied division of decimals by method A, while 69 had studied it by method B. The results show that the students taught by method A had the very low accuracy of 66% in placing the points, while those taught by method B had the very excellent accuracy of 99%. If future investigations confirm these results, method B should be generally adopted.

(8) Speed vs. Accuracy. Thorndike investigated the relation between speed and accuracy in simple addition. Six hundred and seventy-one students were tested apparently on two different occasions in a class experiment in adding columns of nine digits. The subjects were then grouped according to speed as shown by the following table:

TABLE 125. After Thorndike ('15)

Number of Individuals in Group	Number of Additions per 100 Seconds (Counting the Time of Writing the Answer Equal to One Addition's Time)		Approximate Number of Errors per 1000 Additions, i. e., Wrong Answers per 100 Ten-Digit Additions	
	Early Test	Late Test	Early Test	Late Test
65	150	162	7.0	3.8
108	108	120	9.1	6.5
86	88	99	10.3	6.7
115	75	87	12.0	8.3
109	64	75	12.7	9.0
103	55	66	12.6	9.3
65	46	58	14.4	10.5
20	37	46	17.5	14.4

With a decrease in the rate of additions there is a steady increase in the number of errors per 1,000 additions. Thorndike concludes "that the sort of individual who is quick in adding is more accurate also than the one who is slow."

(9) Limits of Attainment. Since there is general agreement that the fundamental number combinations should become automatic association processes, it is pertinent to ask, how high a degree of skill should be developed in pupils? This question is similar to the one discussed in connection with quality of handwriting. It is reasonable to maintain that it is probably uneconomical to attempt to develop a degree of speed beyond a certain point.

How great speed is practically necessary or worth while? Obviously the school may devote relatively too much attention to the development of speed in the four fundamental operations at too great a cost of time. It would, therefore, be important to know what degree of proficiency is needed for the practical affairs of life. An investigation of the problem is needed.

(10) Errors. The detection and classification of errors and the discovery of the frequency with which they occur are highly useful facts in any school subject because such information will help to make instruction specific. It will indicate the particular points at which drill should be directed.

Howell made an analysis of the mistakes in division occurring in the Courtis tests as applied to the pupils in his school. He found the following rubrics of errors:

"1. Making the quotient figure the same as the divisor,
(a) When a difference of only one exists between the divisor and quotient;
(b) When the quotient is commonly used as the divisor of the given dividend.

"2. Making some factor (other than the divisor), commonly used as the divisor of a given dividend, the quotient figure.

"3. When dividing a digit by itself, making the quotient figure *the same.*

"4. When dividing a digit by itself, making the quotient figure *zero.*

"5. When dividing by one, making the quotient one.

"6. When the dividend is zero, making the quotient the same as the divisor.

"7. Pupils whose associations are as yet feeble or become so through fatigue or distraction are commonly observed to resort to running up the table.

"They frequently miss count and get a quotient figure one removed (say) from the right one.

"8. Making one of the quotient figures the quotient.

"9. Substituting multiplication for division.

"10. Unclassified."

The frequency with which the different types of errors occurred is shown in Table 126.

TABLE 126. After Howell

Showing the number of mistakes in the division tables falling into each class

Grade	Kinds of Mistakes										Total Mistakes
	1 a & b	2	3	4	5	6	7	8	9	10	
3	16	12	2	14	1	2	26	6	2	43	124
4	13	7	2	20	4	4	23	3	8	49	133
5	5	7	3	23	3	1	14	0	2	4	62
6	4	8	0	32	16	4	16	2	0	17	99
7	10	6	1	27	4	2	7	2	0	19	78
8	9	9	4	65	11	11	5	3	0	3	120
Total	57	49	12	181	39	24	91	16	12	135	616

From this table it appears that certain types of error occur much more frequently than others. For example, Error No. 4, "when dividing a digit by itself, making the quotient figure zero," occurs 181 times; whereas Error Nos. 3 and 9 occur only 12 times each. More extensive studies of this sort are much needed. Learning in any school subject is apt to be more economical the more specifically the learners' attention may be directed to the particular processes to be exercised.

A. S. Gist tabulated the errors in subtraction, multiplication, and division of 812 arithmetic test papers from six schools in Seattle. Table 127 shows the percentage of different types of errors separately for each of the three processes and for each of the grades from 4 to 8. It is noticeable that the greatest difficulty was presented in subtraction by borrowing, in multiplication by the tables and in division by the remainder. For the most part the proportions of the various errors are fairly constant from year to year.

Besides the prevalence of the general type of errors noted above, the relative difficulty of the various elementary combinations is a matter of much practical importance. C. L. Phelps attempted to determine this for the addition combinations by finding the percentage of errors made on each of the 55 addition combinations of the Courtis tests in 5,950 papers made by repeatedly testing 238 eighth-grade children. These results are shown in detail in Table 128. There is a clear relation between the per cent of errors and the size of the total resulting from the combination. The writer computed the correlation between the two by Spearman's rank method and found a coefficient of .57. The relation is shown in detail by the curve of Figure 76. For the most part the odd and especially the prime numbers are more difficult than the even numbers adjacent to them. While there is seen to be a fairly steady rise up to 10, there is an abrupt increase at the beginning of the teens which continues as far as the investigation extends.

TABLE 127

The per cent of each type of error in the examples of subtraction, multiplication, and division respectively from the 812 papers from six schools in Seattle. (After Gist ('17).)

Grade	4th	5th	6th	7th	8th
Subtraction:					
Borrowing	54	56	52	51	55
Combinations	36	38	45	44	41
Omissions	2	1	2	3	1
Reversions	1	2	½	0	0
7 — 0, 0, etc.	5	3	½	0	0
Left-hand digit	0	0	0	0	2

TABLE 127—Continued

	4TH	5TH	6TH	7TH	8TH
Multiplication:					
Tables	79	73	73	77	75
Addition	18	20	22	19	20
Cipher in multiplier	1.5	6	5	4	5
Division:					
Remainder too large	34	39	27	19	10
Multiplication	22	15	19	37	33
Subtraction	11	14	18	25	23
Last remainder 0, and 0 in dividend	7	15	19	7	11
Multiplicand larger than dividend	7	4	1	1	1
Failure to bring down all of dividend	7	4	3	0	6
Failure to bring down correct digit	2	1	4	4	6
Failure to place all of quotient in quo.	7	1	1	3	3
Cipher in quotient, as 908—98	3	7	8	4	7

A somewhat similar investigation was carried out by Holloway. He tabulated the number of errors made on each of the addition combinations by 1,065 third-grade children. His results are given in Table 128 in the order of the number of errors, parallel to those of Phelps. While there is considerable agreement between the two studies as to the relative difficulty of the various combinations, there are also rather striking disagreements. How much these differences are due to the stress which had been put upon the various combinations in the teaching of the children cannot be determined. The latter defect is much less likely in Holloway's results because of his much wider range of subjects.

Holloway also tabulated the errors in the multiplication combinations in the test papers of 1,215 third-grade children. They are given in detail in Table 129.

TABLE 128

Table showing the relative difficulty of the various addition combinations for the third and eighth grades respectively

Combinations	Number of Errors by 1,065 3rd Grade Subjects.—Holloway	Per Cent of Errors in 5,950 Papers from 238 8th Grade Children.—Phelps	Combinations	Holloway	Phelps
9 + 8	95	2.44	8 + 1	19	.62
9 + 7	90	3.32	3 + 1	19	.45
9 + 6	82	2.60	4 + 3	18	.67
8 + 7	69	2.30	3 + 2	17	.54
8 + 5	68	3.10	6 + 1	17	2.22
8 + 6	66	1.04	1 + 1	17	.27
7 + 5	56	2.25	4 + 2	16	.50
9 + 4	51	....	9 + 2	15	1.16
7 + 6	50	1.56	5 + 1	15	.72
9 + 5	49	2.50	4 + 1	15	1 30
7 + 4	48	1.95	5 + 2	13	.86
9 + 3	43	2.55	9 + 1	13	1 09
8 + 3	41	1.94	8 + 2	13	.88
8 + 8	37	1.98	5 + 5	9	.07
8 + 4	37	2.38	2 + 2	9	.37
7 + 3	37	2.02	4 + 4	8	.12
6 + 4	34	.69	3 + 3	8	1.46
6 + 5	32	2.27	0 + 8		.62
9 + 9	29	.30	0 + 5		.52
5 + 3	26	1.31	0 + 0		.39
7 + 2	24	1.32	0 + 3		.28
2 + 1	21	.64	0 + 1		.25
7 + 7	20	.14	0 + 4		.24
6 + 6	20	.35	0 + 7		.24
5 + 4	20	.72	0 + 9		.15
6 + 3	20	1.19	0 + 2		.14
7 + 1	20	2.42	0 + 6		.05
6 + 2	19	.71			

TABLE 129

Table showing the order of difficulty as determined by the number of errors of 1,215 third grade children at end of year. After Phelps.

11 × 11	735	11 × 9	181	5 × 4	55
12 × 11	655	7 × 5	181	6 × 2	50
11 × 10	638	9 × 3	169	5 × 3	46
12 × 10	542	9 × 5	168	11 × 2	46
12 × 8	460	11 × 8	167	1 × 1	41
9 × 7	455	8 × 3	151	9 × 2	39
12 × 7	438	11 × 6	144	10 × 3	38
8 × 7	435	6 × 5	138	7 × 2	38
12 × 12	425	11 × 7	137	5 × 5	34
9 × 8	422	8 × 5	137	4 × 2	32
12 × 9	417	6 × 4	133	10 × 4	31
9 × 6	390	6 × 6	129	10 × 2	31
8 × 8	361	11 × 5	113	11 × 1	31
12 × 6	361	6 × 3	102	4 × 1	31
8 × 63	42	11 × 3	99	3 × 1	28
9 × 4	292	10 × 9	94	5 × 2	26
7 × 6	285	10 × 7	86	3 × 3	25
12 × 5	271	10 × 8	85	9 × 1	22
7 × 7	268	12 × 2	81	3 × 2	21
9 × 9	263	10 × 6	79	9 × 1	21
12 × 4	250	4 × 4	78	6 × 1	21
10 × 10	241	4 × 3	76	12 × 1	20
8 × 4	235	7 × 3	71	5 × 1	20
7 × 4	192	10 × 5	58	2 × 1	20
12 × 3	183	8 × 2	58	2 × 2	18
				8 × 1	18
				10 × 1	12

CHAPTER XXI

HISTORY

Psychological Processes in Learning History

It is more difficult, at least more uncertain, to make an analysis of the psychological processes concerned in the learning of history than it is in case of most of the subjects treated thus far, for the reason that teachers are not as fully agreed as to what is to be learned in history. In general there are two extreme views: One would hold that the learning of history means the learning of the main facts—names, dates and events—of the human race; while the other would hold that history means the acquisition of ability to interpret the significance of human events. In their most extreme forms the two views would be a memorizing of isolated facts versus an interpretation of facts with little emphasis on facts. Probably no one holds either of these extreme positions and the distinction is useful only for analytical purposes since the psychological processes involved in learning history would be quite different in these two modes of approaching the subject. Those who stress the interpretational aspect of history would stress the conception that the purpose of history in the public schools is training for citizenship or the development of patriotism.

Let us assume for our present analysis that the obvious aim of what is to be learned or acquired in history is facts of the events of human beings and the connections among these facts—a view to which probably every one could agree. What are the psychological functions concerned in acquiring and connecting historical facts? Let us take as a concrete instance a given historical event and see what mental operations are necessary to grasp, interpret and remember it. Let us take the statement that Columbus discovered America in 1492. The psychological processes involved or assumed in learning and grasping this statement would be substantially as follows: To begin with, it would involve all the steps enumerated in the analysis of the reading process, or of the process involved in understanding spoken language when the facts are heard instead of read, since practically all history is learned through reading. The chief difference would be in the emphasis and elaboration of some of the steps. Taking the factors as enumerated in Chapter XVI, the main difference between ordinary reading and learning history would be in step (6): "The establishing or arousal of association processes whereby the incoming impulses are interpreted." These specific processes of association and interpretation over and above ordinary reading, necessary for grasping a historical statement are as follows:

(1) A mental picturing or conceiving of the persons, actions, localities and objects, concerned in the event.

(2) A mental picturing or conceiving of points and locations in time.

(3) Processes (1) and (2) applied to events preceding the one in question.

(4) Processes (1) and (2) applied to events succeeding the one in question.

(5) A judgment concerning the internal motives of the persons and the external conditions that led to the event.

(6) A judgment concerning the effect of the event upon the motives of the persons and environmental conditions involved in succeeding events.

(7) A remembering of the mental processes in steps (1) to (6), in so far as a permanent memory of them is considered important.

Steps (1) to (4) use primarily the imagination, steps (5) and (6) judgment and reasoning, and step (7) memory.

Thus the grasping, interpreting, and remembering of the statement that Columbus discovered America in 1492 would mean, (1) An imagining or conceiving of Columbus as a person, his associates, ships, water, the voyage and landing on new soil; (2) an imagining or conceiving of the time so as to give a notion of how long ago 1492 was; (3) and (4) a similar procedure with events coming before and after this particular one, such as Columbus's interview at the court of Queen Isabella of Spain, the need for another route to India, succeeding voyages, settlement of the new country, and the like; (5) and (6) judgments concerning the effect of the preceding events in bringing about the particular event under consideration, and judgments concerning the effect of the latter in bringing about later events; and (7) a repetition of the learning of the fact with its connections to fix it in mind.

Measurement of Attainment in History

The difficulties that beset any endeavor to devise an objective and generally acceptable method of measuring attainment in history are very great for the reason that historians differ very widely in the selection, emphasis and interpretation of facts and in the manner of stating the facts; and also for the reason that teachers as well as texts differ in the relative emphasis on the learning and remembering of facts as compared with their interpretation. The author has prepared a plan with the aim of meeting these difficulties as far as possible so as to secure a test that could be used fairly wherever American history is taught; it is necessary to exercise much care in selecting the right material for testing purposes. The scheme finally carried out was as follows: Five widely used text-books in American History were carefully compared and all facts and interpretations given in all five were selected and formulated into statements or sentences. This gave a total of 278 statements—a remarkably small body of facts common to five texts. These statements were then made into a mutilated text or completion test. Certain important words or phrases were omitted which are to be supplied by the pupils doing the test. The entire 278 statements would be too long as a single test. Hence they were split up into four parallel sets, each containing 69 or 70 statements, by taking for the first set, statements numbers (1), (5), etc.; for the second set, numbers (2), (6), etc.; these four tests may be used interchangeably at different times in testing a class. Direct comparisons and measurements of progress can thereby be made. The score of a pupil is the number of omitted parts correctly supplied. The following statements serve to illustrate the nature of the resulting test. They are the first ten statements of the first test.

1. ——————— discovered America in 1492.
2. John Cabot exploring for the ——————— in 1497 landed on the ——————— coast and claimed the country for ———————.
3. ——————— sailed around the globe in 1519–1521.
4. ——————— discovered the Mississippi River in 1541.
5. Two expeditions sent out by ——————— to settle Virginia in 1585 and 1587 respectively, failed.
6. ——————— was governor of Virginia after Delaware left.

7. ——————in service of the Dutch East India Company, explored the——————river in 1609.

8. ——————was Governor of the Dominion of New England, which was composed of (1)——————(2)——————and (3)——————.

9. John Winthrop came to America in 1630 and settled——————.

10. New Hampshire was founded in——————.

The following are the average scores for the ends of the different grades obtained from approximately 2,000 pupils:

Grade	6	7	8	H. S.
Scores	7	20	38	38

Individual differences and overlapping of successive grades as shown in Figure 22 are enormously wide. In a certain eighth grade composed of thirty-six pupils, the best pupil made a score of 102 and the poorest a score of 4. The differences among various schools are indicated in Table 130. Very wide differences exist, which are probably due chiefly to differences in methods of teaching. Thus the best eighth grade made a score of 66 and the lowest one a score of 19. Even in the same school system the differences among schools may be very wide. For example, in city N, the best eighth grade averaged 52 and the poorest 19. Another striking observation is the fact that the average attainment of the high school pupils is no better than that of the eighth grade. Apparently the pupils relearn in the high-school course in American history about as much as they forgot during the intervening two or three years since they left the eighth grade. This does not necessarily mean that American history in the high school is useless since the relearning will help to guard against further loss. More extensive tests are needed on this point.

Striking sex differences in knowledge of history have been revealed by such a test as the one here described. The median scores for boys and girls in the writer's test were as follows:

	Boys		Girls	
	Number	Median Score	Number	Median Score
High School	47	41	73	36
8th grade	288	45	352	31
7th grade	94	24	101	17

Bell and McCollum applied a history test to 1,500 pupils and found similar differences. The boys in the elementary schools did 28% better and in the high schools 31% better than the girls. This superiority on the part of the boys may be due to their greater interest in battles which in turn may be due to the greater strength of the fighting instinct.

TABLE 130

Scores in the writer's American History Test, Series A

Grade	7	8	H. S.
City A	—	66	35
" B	—	30	—
" C	—	32	48
" D, School 1	—	22	—
" " 2	—	59	—
" " 3	—	—	42
" E	17	33	—
" F	—	40	—
" G, School 1	22	—	—
" " 2	17	65	—
" H " 1	17	40	—
" " 2	19	48	—
" I " 1	20	—	—
" " 2	—	33	—
" " 3	—	—	39
" J	—	—	41
" K	15	36	—
" L	—	32	—
" M	—	29	—
" N School 1	—	33	—
" " 2	—	19	—
" " 3	—	52	—
" " 4	17	50	—
" " 5	—	29	—

ECONOMIC METHODS IN THE LEARNING AND TEACHING OF HISTORY

Experimental work that has been done up to the present time in the psychology and pedagogy of school subjects has been confined almost entirely to the subjects thus far considered. Yet the problems and factors entering into such a subject as history are exceedingly intricate and as much worth while and for the most part as capable of experimental determination as most of the problems in the other subjects. The discussion of factors and conditions affecting the most economic procedure in history will, therefore, have to be limited to a few suggestions and beginnings in experimental work. The problem for the future will consist in determining the factors which promote or retard the elements enumerated in the first section of this chapter. Economic procedure in learning history resolves itself into discovering the most favorable means of, and a measurement of their actual effects in, grasping, imagining, judging and remembering the important events of the human race.

(1) By what means may the imagining or conceiving of a given event be brought about most effectively? Numerous devices are employed to assist the imagination, such as pictures, dramatizations, pageants, etc. These are probably used with profit but no one has ever determined to what extent they actually contribute, or whether they contribute at all to the better understanding of the event, or how much time devoted to dramatization, for example, is worth the returns it may bring. Experimental work should obviously be undertaken. Dramas and pageants may be easily overdone and may often deal with unimportant phases of the persons or events concerned.

(2) By what method may the imagining or conceiving of a given event in time with respect to other events be accomplished? Orientation in time is a very complex psychological process and probably develops rather gradually through the years of a child's experience. It probably develops from the immediate perceptions of changes in the child's environment to the gradual extension to longer historical periods which are not directly perceived but are thought of in symbolic form. Thus the writer pictures different periods and points in history in spatial terms by imagining a horizontal line about three feet in length extending from a point, which represents the present, toward the left, that is back to the past. The discovery of America is located about four inches to the left from the beginning point, the birth of Christ about a foot and a half to the left, and so forth for other approximate locations in time. As a mere opinion, the writer believes that it would be advantageous to introduce the pupil to the study of history by giving a bird's-eye view over long stretches of time, as this would probably aid the imagination in conceiving time. Historians usually object rather strenuously to this method of introducing or teaching history. Psychologically, it would seem easier to imagine long periods of time and the relative location of events in them by viewing all history pretty much at a glance, and by giving then

more and more detailed consideration to each period. But this again is a problem requiring experimental determination.

(3) How may judgments about the personal motives of historical figures and causal relations among events be best developed? Nobody knows. All we can say is to encourage the making of such judgments and interpretations according to the best insight of the pupil and then to check them up with those of competent historians.

(4) What are the most effective methods of remembering historical events? In spite of the extensive experimental work in the field of memory, there is very little in the way of concrete advice that can be given to a pupil to assist him at this point. The following, partly general and partly specific, suggestions may be given. These have been corroborated by experimental data and have been stated in Chapter XII on How to Study, and in Chapter XVI on Reading, and will, therefore, be only mentioned here.

a. Thoroughly understand the facts you wish to remember.

b. Systematize the facts to be remembered.

c. Look for the essentials.

d. Recall, after every paragraph or two, the essential ideas read.

e. At longer intervals, re-think or review the essential ideas again.

f. Develop your own special means, associative links, or schemes for remembering certain facts. Systems of memory such as that developed by Loisette for remembering, for example, the names of the Presidents of the United States, consist in establishing certain associations of similarity in sound between certain parts of the successive names, as shown in the following illustration:

George WashingTON JOHN Adams	In.	"Ton" and "John" make a fairly good In. by sound.
JOHN Adams THOMas Jefferson	In.	"John" and "Thom" (the "h" is silent in both names) make an IN. by sound, imperfect but adequate if *noticed*.
Thomas JefferSON James MadiSON	In.	Both names terminating with the same syllable, "son," makes a clear case of In. by sound and spelling.
JAMES Madison JAMES Monroe	In.	This pair of names furnishes an example of perfect In. by sound and spelling in the Christian names.

James MONroe JOHN Q. Adams	In.	"Mon" and "John" give us a good In. by sound.
JOHN Q. Adams Andrew JACKson Etc.	In.	"Jack" is a nickname for John—a case of Synonymous In. (Loisette, p. 26.)

The main objection to such a plan is its artificialty. The chief advantage is that it does draw specific attention to the facts to be remembered; but everyone can develop his own special links or clues for retaining facts with which he has difficulty. These are likely to be more natural, more serviceable and more permanent than artificial ones forced upon the learner from the outside. The main point is that each one should attempt to establish such clues which will usually result in discovering useful associative links and at the same time force attention upon the facts to be retained.

(5) Another very important problem in the economy of learning history is the question of essential material. What should the child really be expected to master? What facts, names and dates should he actually learn? What interpretations should he be led to make and acquire? In general there has been a distinct shift from regarding history as a chronicle of wars to regarding history as a tracing of the development of political, industrial and social institutions. In recent years, various committees have been at work to decide upon a body of minimum essentials. Thus the committees of Iowa and Minnesota have made the following suggestions as summarized by Betts ('17):

"Wars. Limit the study of wars to their remote and immediate causes; their general geography; resources and problems of nations involved; general plan of military operations; a few critical battles; important leaders; what the war settled, and the after effects; cost in men and treasure. This plan will reduce the war phase of history study by more than half.

"Eliminate the detailed study of battles except: Battle of Quebec; Lexington and Concord; Bunker Hill; Saratoga; Yorktown; Lake Erie; Merrimac and Monitor; Gettysburg; Vicksburg; Manila.

"Dates. Limit the memorizing of dates to events of central importance like the following: 1492, the discovery of America; 1607, settlement of Jamestown; 1619, slavery introduced; 1620, Pilgrims land at Plymouth; 1643, confederation of colonies; 1775, Lexington, Concord and Bunker Hill; 1776, Declaration of Independence; 1781, Cornwallis surrenders; 1789, First Congress; 1793, Whitney's cotton gin; 1803, Louisiana Purchase; 1807, Fulton's steamboat; 1812, war with England; 1820, Missouri Compromise; 1823, Monroe Doctrine; 1826, first railroad; 1844, first telegraph; 1846, sewing-machine invented; 1845, first reaper; 1846–1848, Mexican War; 1861, secession and Civil War; 1863, Emancipation Proclamation; Gettysburg, Vicksburg; 1866, Atlantic cable; 1876, first telephone, 1878, electric light invented; 1898, war with Spain; 1903, first wireless across Atlantic; 1914, world war in Europe.

"Other omissions. Detailed provisions of various tariff acts (but the meaning of tariff should be understood); details of political campaigns except Jefferson's, Jackson's, Lincoln's and any current campaign in progress; critical study of political party principles (but give broad distinctions between chief rival parties); financial panics except those of 1837, 1873, 1893." (Pp. 271 and 272.)

Bagley ('15) has been engaged on working out a possible scientific plan for determining the relative emphasis upon various portions of history by discovering the frequency of reference to persons and events made in current magazines and newspapers. He concludes that such a study is suggestive but doubts whether it may serve as a final criterion for determining the amount of time and emphasis to be given to various phases of history.

Bagley and Rugg ('16) made a study of twenty-three text-books, published between 1865 and 1912, by comparing the amount of space given to various topics in each book, by determining the shift in emphasis in the course of this period of time as measured by the space given to different topics, and by listing the topics and names included in all, or in a certain fraction of these texts.

"(1) In so far as can be determined from the materials presented in the text-books, elementary American history as taught in the 7th and 8th grades has been and still is predominantly political and military history.

"(2) Within the past fifty years, the emphasis upon military affairs as measured by the proportion of space devoted to wars has declined. In general, battles and campaigns are treated less in detail than was formerly the rule, while proportionately more space is devoted to the causes and the results of the wars. The lessening emphasis upon details of the wars is first noticed in some of the text-books published between 1881 and 1888, and the tendency has been general and decided since that time.

"(3) The later books give a perceptibly heavier emphasis to the facts of economic and industrial development than do the earlier books, although political development still constitutes the essential core of elementary historical instruction.

"(4) As regards the treatment of specific eras or epochs, the principal increases in emphasis are to be noted in connection with: (a) the period

1783–1812 (especially in the treatment of the so-called 'critical period' between the close of the Revolution and the adoption of the Constitution); (b) the non-military affairs of the period 1812–1861; and (c) European events preceding and during the periods of discovery, exploration, and settlement.

"(8) Numerous changes have taken place in the construction of elementary text-books in history during the past fifty years. The more important of these are: (a) a movement toward a simpler 'style' with larger emphasis upon clear statements of causal relationships: (b) the introduction and development of the 'problem' as a method of teaching history, and a consequent encouragement of 'judgment' as contrasted with rote memory,—of rational as contrasted with verbatim mastery; (c) a marked decline in the employment of imaginative pictures as illustrations and an increase in the use of pictures that represent sincere attempts to portray actual conditions; (d) a marked decline in the use of anecdotal materials; (e) a larger and wider use of maps." (Pp. 56 and 57.)

Horn ('17) conducted an investigation after the manner of Bagley's magazine-newspaper method by checking through twenty-seven recent books on current industrial, political and social problems, in order to ascertain the facts, persons and dates referred to and the frequency of reference. His general impression of this inquiry is stated thus:

"This investigation has not attempted to answer the question as to the complete content of the course of study in history. Neither does it assert that the purpose of history is to throw light on modern social problems, or that this is even one of the chief purposes of studying history. Without regard to what the aims of teaching history are, this investigation has been carried on to examine into the implications of one particular assertion: namely, that history should render pupils more intelligent with regard to modern conditions, problems and activities. If one assumes (1) that this is the function of history, (2) that the method of research here followed is satisfactory, and (3) that sufficient data have been collected, then there seems to be no escape from the conclusion that the present elementary and high-school courses of study in history are in very serious need of reconstruction." (Horn, '17, p. 171.)

CHAPTER XXII

MARKS AS MEASURES OF SCHOOL WORK

Importance of Marks. In order to determine the fruitfulness or wastefulness of methods of learning and teaching school subjects, it is necessary to evaluate the achievements of pupils as accurately as possible. Furthermore, the successful operation of a school demands an accounting of the work of its pupils.

Marks have been the universal measures of school work. So many problems in the management of a school—credit, failure, promotion, retardation, elimination, graduation, honors, recommendations for positions, indeed the entire scholastic machinery of a school—hinge upon the assignment of marks that it is highly imperative to examine in detail the value, accuracy and reliability of marks as well as to ascertain the possibility of some sort of standardization of marks.

Variations among Teachers and Schools in the Distribution of Marks. The manner in which marks are distributed to pupils varies enormously from teacher to teacher and from school to school. No one realized the seriousness of the situation until specific tabulations and comparisons were made.

Meyer published the distribution of the marks assigned by 40 different professors at the University of Missouri to their students during a period of five years as exhibited in Table 131.

TABLE 131. After Meyer ('08)

Distribution of the marks of 40 teachers in the University of Missouri for a period of five years. The numbers are the percentages receiving the various grades.

TEACHERS	A	B	C	F	TOTAL NO. OF STUDENTS
Philosophy	55	33	10	2	623
Latin I	52	42	6	0	130
Sociology	52	30	13	5	958
Mathematics I	40	31	16	13	208
Economics	39	37	19	5	461
Greek	39	26	24	11	287
Latin II	36	40	19	5	577
French	36	29	25	10	295
Political Science	34	30	27	9	592
Mathematics II	32	29	23	15	145
German I	30	29	20	11	586
Psychology I	30	36	24	10	907
German II	26	38	25	11	941
Elocution	20	61	19	0	917
Geology	22	48	22	8	293
History I	14	53	27	6	779
Zoölogy I	21	45	28	6	479
Psychology II	19	47	29	5	238
History of Art	25	40	30	5	685
Bacteriology	20	45	31	4	263
Freehand Drawing	18	47	25	10	506
Chemistry I	23	40	31	6	205
English I	21	41	30	8	964
Astronomy	13	49	33	5	225
History II	11	51	33	5	806
Zoölogy II	24	37	31	8	250
German III	22	37	28	13	441
Chemistry II	9	48	43	0	21
Education	18	38	35	9	266
Mathematics III	19	36	26	19	182
Mathematics IV	25	29	36	10	380
Physiology	20	33	40	7	426
Anatomy	19	34	36	11	544
Mathematics V	16	34	35	15	209
Engineering I	13	36	42	9	813
Mechanical Drawing	18	29	41	12	538
Mechanics	18	26	42	14	495
Engineering II	16	26	46	12	826
Chemistry III	1	11	60	28	1 903
English II	9	28	35	28	1,098

Similar tabulations have been published for Harvard University by Foster, Table 133; for the University of Wisconsin by Dearborn, Table 134; and for Cornell University by Finkelstein. Finkelstein ('13) has shown the effect of the personal equation in the marks assigned to the same students in a year course which was in charge of one teacher in the first semester and of another in the second semester.

TABLE 132. After Finkelstein

	No. of Students	Per Cent Receiving the Various Marks						
		0–39	40–44	45–49	50–54	55–59	60–64	65–69
1st semester ...	263	.4	.4	2.3	4 5	1.9	13 7	11.8
2nd semester ..	257	—	—	1.2	.8	1.5	12.1	12.8

	No. of Students	70–74	75–79	80–84	85–89	90–94	95–100	Exempt
1st semester, con...	263	16.7	15.6	20.2	6.4	5.7	.4	12.5
2nd semester, con...	257	10.5	13.6	9 8	33.9	3.8	—	37.7

The instructor in the first semester exempted from the final examination 12.5% of the class, while the instructor in the second semester exempted 37.7%. The latter obviously graded very much higher than the former.

TABLE 133. After Foster ('11).

Harvard College. Distribution of 8,969 grades. Elementary Courses

GROUP I	A%	B%	C%	D%	E%	Abs %	TOTAL
Astronomy	16	13	45	19	6	1	69
	10	17	48	17	7	2	130
Botany	11	28	38	14	2	7	183
	4	32	44	13	½	6	219
Chemistry	6	26	45	12	9	2	334
	8	19	45	17	11	0	319
Economics	10	18	37	25	7	3	531
	7	19	43	21	7	3	436
Engineering	11	15	31	28	12	3	114
Engineering	10	13	27	21	21	9	121
							139
							129
English	1	13	52	28	3	3	603
	1	11	51	32	3	3	564
Fine Arts	2	33	45	10	2	9	58
	6	27	67	0	0	0	49
French	11	25	35	21	4	4	156
	12	19	36	19	10	4	145
Geology	5	26	45	20	3	2	489
	5	25	33	28	2	7	85
Geology	2	28	48	10	7	5	122
	4	20	43	24	7	2	108
German	7	21	31	26	11	4	259
	6	14	32	27	17	2	293
Government	6	16	39	28	8	3	356
	9	23	37	21	7	2	419
Greek	35	28	21	13	1	3	72
	15	36	34	7	5	3	61
History	7	20	44	21	5	2	347
	7	24	42	20	6	2	380
Hygiene	18	29	33	18	1	0	87
	8	23	48	14	4	3	139
Latin	17	25	41	10	7	0	143
	15	27	41	5	10	2	128
Mathematics	18	24	18	31	11	0	85
	14	22	31	23	11	0	95
Philosophy	7	31	41	15	2	5	229
	7	23	61	8	0	½	215
Spanish	10	24	43	16	4	3	106
	7	13	38	33	8	2	119
Zoölogy	2	13	48	30	5	1	149
							184
Averages	7	20	42	21	7	3	213

TABLE 134

Percentages of grades assigned by 45 individual instructors in the University of Wisconsin to freshmen and sophomores. After Dearborn ('10)

	Ex.	G	F	P	X	No. of Cases
History:						
1	4.9	26.2	32.8	26.2	9.8	183
2	9.8	52.9	31.6	3.11	2.5	193
3	3.4	22	38.9	26.7	8.9	558
4	7.4	25.2	37.4	23.4	6.5	107
5	16.7	52.4	28.6	2.3	0	42
6	9.1	39.4	27.3	18.2	6.0	33
7	6.3	27.4	30.8	23.6	11.8	237
English:						
8	19 3	22 6	19.3	38.7	0	31
9	12 5	30	37.5	12.5	7.5	40
10	6.4	45	31.2	14.7	2.7	109
11	1.9	33.2	45.5	12.9	6.4	202
12	7.4	30 6	38.0	19.0	4.9	121
13	6 3	33 7	40	14.7	5.2	95
14	4.0	38 8	32.6	14.3	10.2	49
15	16.7	22.8	36.0	17.5	7.0	114
16	7 3	39.6	37.5	13.5	2.0	96
17	16.9	50.6	19.5	11.7	1.3	77
18	11 4	42.9	27.2	14.3	4.2	70
19	8.9	53.9	21.8	7.6	7.6	78
20	6.1	25.5	23.5	37.8	7.1	98
Mathematics:						
21	16.6	27.8	28.5	14.2	12 9	302
22	24 1	25	15.7	19.4	15.7	108
23	12.1	17.9	19.3	27.4	23.3	223
French:						
24	21.2	43.8	22 5	7.5	5	80
25	17.5	43.8	27 5	7.5	3.7	80
26	22.3	35.7	21.4	12.5	8.0	112
27	15.5	32	27.2	19.4	5.8	103
28	14.5	24.2	24.2	27.4	9.6	62
29	23.9	35.4	24.6	13.8	2.3	130
Physics and Chemistry:						
30	27.9	37.8	21.1	10.8	2.4	204
31	21.1	35	24.6	15.6	3.8	289
Latin:						
32	16.1	46	20.7	9.2	8.0	87
33	11.7	46.2	26	16	0	119
34	26.1	47.6	15.9	7.4	2.8	107

German:

35	26.3	34.2	21.9	12.3	5.2	114
36	12.0	49.1	24.1	11.1	3.7	108
37	34.3	40.3	19.4	2.9	2.9	67
38	11.4	34.4	27.9	22.9	3.2	61
39	17.3	37.3	29.3	6.6	9.3	75
40	17.9	35	26	17.1	4.0	123
41	14.7	29.5	33.5	18.9	3.1	95
42	12.3	27.4	27.4	21.9	10.9	73
43	29.0	30.1	22.6	14	4.3	93
44	21.6	42.0	21.6	12.5	2.27	88
45	22.4	39.6	20.7	15.5	1.7	58

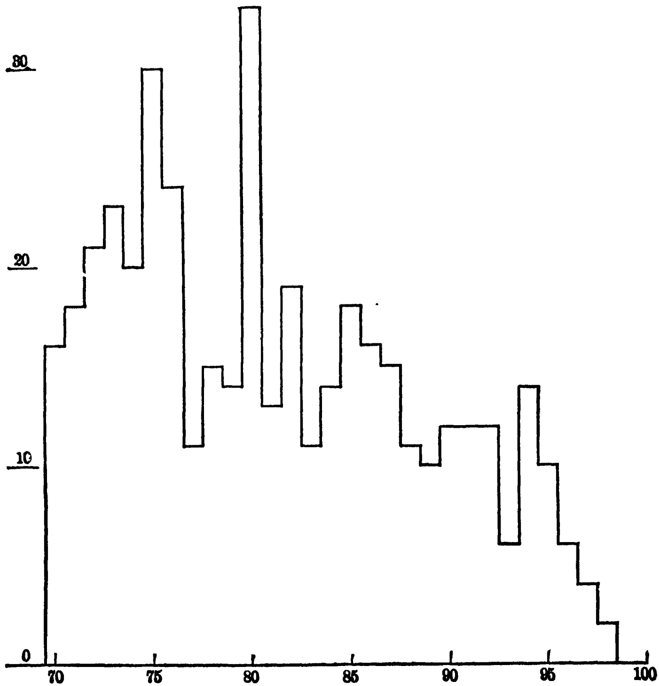

FIG. 85.—Distribution of all grades in a high school. After Gray ('13, p. 66).

These tables agree in showing extremely wide differences among teachers in the manner of giving marks. In the tabulation for the University of Missouri, one professor assigned the grade A to 55% of his students, and the grade F to only 2%, while another professor assigned the grade A to only 1% of his students and the grade F to 28%. At Harvard, one professor gave the grade A to 35% and the grade E to 1%, while another professor gave A to 1% and E to 32% of his students.

The situation in high schools is substantially the same. Gray tabulated the marks assigned by all the teachers in eight high

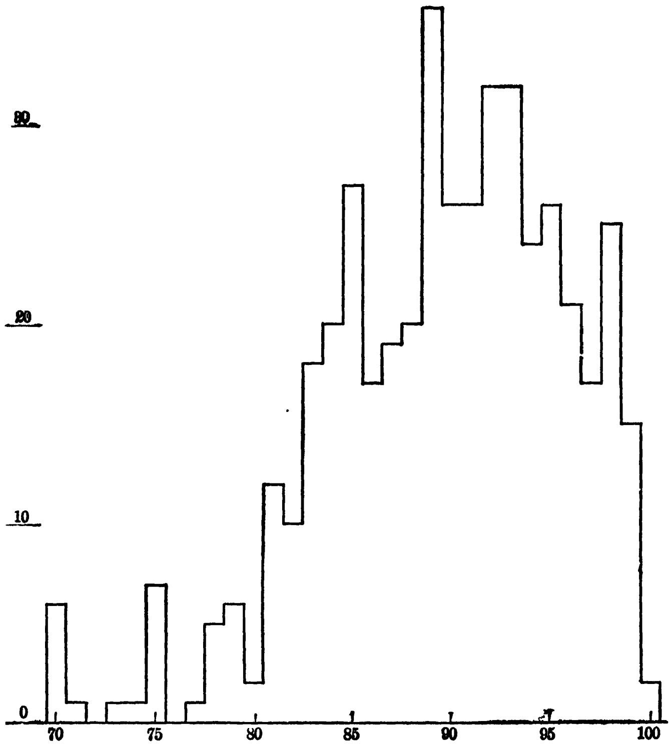

FIG. 86.—Distribution of all grades in another high school. After Gray ('13, p. 67).

schools. The distributions of two of these schools are shown in Figures 85 and 86. The one grades high and the other grades low. In one school the great mass of the pupils receive 85 to 100; in the other they receive 85 to 70.

Table 135 shows the distribution of the grades by the different teachers in a high school of about 150 pupils.

TABLE 135

Distribution of 8,490 grades by the different teachers in a high school during four years. (From a private report by Superintendent J. F. Waddell, Evansville, Wisconsin.)

	−74	75–80	81–86	87–92	93–100
English (1913–1914, 1914–1915)	33%	39%	13%	12%	3%
English (1915–1916)	15	27	27	25	6
Latin and German	5	15	13	33	34
Mathematics (1913–1914)	22	31	18	21	8
Mathematics (1914–1915)	23	27	24	13	13
Mathematics (1915–1916)	5	12	15	28	40
History (1913–1914)	11	25	22	33	9
History (1914–1915; 1915–1916)	10	18	25	30	17
Science (1913–1914)	11	36	25	26	2
Science (1914–1915; 1915–1916)	10	33	24	19	14
Domestic Science	0	12	27	51	10

Distribution of grades for the year after the above tabulation was made known to the teachers. Extreme variations are considerably reduced.

English (1916–1917)	6%	32%	27%	27%	6%
Latin and German (1916–1917)	9	25	22	29	15
Mathematics (1916–1917)	5	20	28	27	20
History (1916–1917)	2	17	37	34	10
Science (1916–1917)	5	26	33	22	14
Domestic Science (1916–1917)	0	10	40	41	9

Variation among Teachers in the Evaluation of the Same School Products. A more direct and crucial method of examining the variations of teachers' marks than the tabulation of the grades as distributed by different teachers is to measure experimentally the differences in the values assigned by different teachers to the same pieces of work.

Starch and Elliott ('12 and '13) made a series of investigations in which two final examination papers in first-year-high-school English were graded by 142 English teachers in as many high schools, one final examination paper in geometry was graded by 118 teachers of mathematics, and one final examination paper in American history was graded by 70 teachers of history. The variations in these marks are shown in Figures 87, 88, 89 and 90. The differences are astounding; the marks for any given paper run practically over the entire range of the percentage scale ordinarily used. The marks of the first English paper run all the way from

64 to 98, of the second English paper from 50 to 98, of the geometry paper from 28 to 92, and of the history paper from 43 to 90.

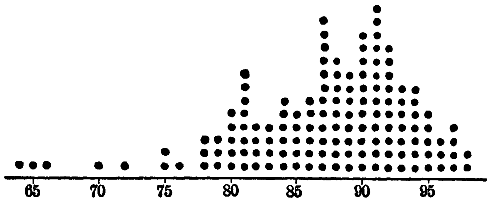

FIG. 87.—Distribution of the marks assigned by 142 English teachers to a final examination paper in high-school freshman English. After Starch and Elliott ('12).

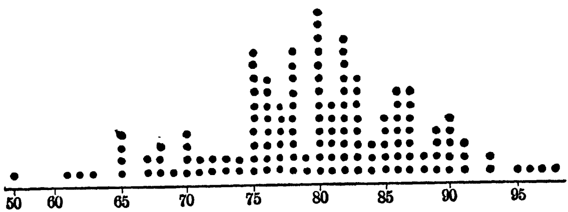

FIG. 88.—Distribution of the works assigned by 142 english teachers to another final examination paper in high-school freshman English. After Starch and Elliott ('12).

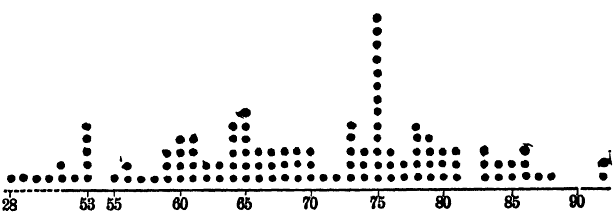

FIG. 89.—Distribution of marks assigned by 114 mathematics teachers to a final examination paper in geometry. After Starch and Elliott ('13).

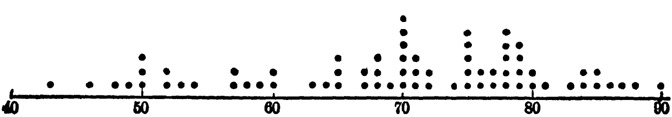

FIG. 90.—Distribution of marks assigned by 70 history teachers to a final examination paper in American history. After Starch and Elliott ('13).

This investigation has established two conclusions: first, that teachers differ enormously in evaluating the same pieces of work in terms of the ordinary percentage scale; and second, that they differ as much in one subject as in another. They disagree as much in evaluating a paper in mathematics as in English or history. Apparently mathematical papers are not marked with mathematical precision any more than any other papers are.

The author made a further investigation by having ten final papers in freshman English in the University of Wisconsin graded by ten instructors of freshman English. The marks are shown in the following table:

TABLE 136. After Starch ('13)

Marks assigned by ten instructors to ten final examination papers in English.

Papers	Instructors 1	2	3	4	5	6	7	8	9	10	Average	Mean Var.	Coefficient of Variability
1	85	86	88	85	75	80	88	87	85	87	84.6	2 8	.034
2	77	80	87	80	62	82	82	87	85	87	80.0	4 6	.057
3	74	78	78	75	69	84	91	83	79	80	79 1	4.4	.056
4	65	65	62	20	26	60	55	68	55	50	52.6	12.3	.233
5	68	82	78	82	64	88	85	86	78	80	79.1	5.7	.070
6	94	87	93	87	83	77	89	88	88	89	87.5	3 2	.036
7	88	90	95	87	79	85	96	91	87	89	88.7	2.6	.029
8	80	84	73	79	72	83	85	91	77	76	80 0	4.6	.058
9	70	70	68	50	44	65	75	81	79	79	68.1	9.1	.118
10	93	92	85	92	81	83	92	89	84	85	87.6	4 0	.045
Av.	79.4	81.4	79.8	73.7	65.5	78.7	83.8	85.1	79.7	80.2	78.7	5.3	.074

The variations shown in this table are practically as large as those found in the previous inquiry. It was thought that the wide range of marks shown in the first study might be due to the fact that the teachers were in different schools. However, Table 136 shows that teachers in the same department differ almost as much. Less extensive results obtained by having various members of a department grade the same paper show that as much variation exists in other subjects as in English.

Causes of Variation. Why do teachers differ so much in estimating the worth of a given product and in the distribution of marks to groups of pupils? Four possible factors may be mentioned: (1) Differences in the standard of severity or leniency in different schools; (2) differences in the standard of severity or leniency of different teachers; (3) differences in credit or penalty assigned by different teachers to any given fact or error in a piece of work; and (4) minuteness of the discrimination between successive steps of merit or quality in a given scale of qualities. How potent is each factor in producing the total variation in evaluating a given paper?

The first thought that occurs in regard to the wide range of marks for the same papers as shown in Figures 87 to 90, was that it must be due to the fact that these teachers were situated in different schools with different standards and ideals. It turns out, however, that factor one is relatively insignificant. If we compare the mean variation of the marks of the ten English papers assigned by ten instructors in the same department with the mean variation of the marks of the two English papers assigned by teachers in different schools we can determine the part played by factor one. The mean variation of the former set of marks is 5.3 and of the latter is 5.4. Hence the mean variation of the marks assigned by teachers in the same department is only 0.1 less than the mean variation of marks assigned by teachers in different schools.

The potency of factor two may be ascertained from the data in Table 136. The general average of all the grades assigned by the ten instructors to the ten papers is 78.7. If we compare the average of all the marks given by any one instructor with the general average we obtain a measure of his particular standard of severity or leniency. Thus instructor 5 graded on the average 13.2 points lower and instructor 8 graded 6.4 points higher than the general average of all the teachers. If now we raise or lower each instructor's grades by as many points as the average of his grades is below or above the general average, we find that the mean variation of these weighted marks is 4.3. This mean variation is only 1.0 point smaller than the mean variation of the original unweighted marks. Hence factor two accounts for a relatively small share of the total variation. Factors three and four must then account for the remaining mean variation of 4.3. The strength of factor four can be determined experimentally by having the same teachers re-grade their own papers without knowledge of their former marks. The author carried out such an experiment and obtained the results exhibited in Table 137.

TABLE 137. After Starch ('13)

	Advanced Psychology Interval 2 Yrs.			Elem. Psychology Interval 2 Weeks			Mathematics Interval 9 Months			Mathematics Interval 9 Months			English Interval 6 Months			German Interval 6 Months			Elem. Psychology Interval 4 Yrs.		
	1st	2nd	Dif.	1st	2nd	Dif.	1st	2nd	Dif.	1st	2nd	Dif.	1st	2nd	Dif.	1st	2nd	Dif.	1st	2nd	Dif.
	85	87	2	85	79	6	36	51	15	56	60	4	70	75	5	79	70	9	70	80	10
	76	80	4	87	83	4	61	67	6	70	73	3	80	86	6	90	77	13	93	91	2
	83	80	3	90	93	3				77	75	2	88	88	0	77.5	73	4.5	82	84	2
	89	90	1	90	92	2	61	67	6	88	90	2	74	76	2	85	81	4	75	82	7
	84	83	1	83	88	5	73	79	6	62	62	0	77	76	1	78	80	2	75	86	11
	93	88	5	78	79	1	81	86	5	89	87	2	85	86	1	70	61	9	78	81	3
	84	75	9	93	89	4	71	63	8	82	80	2	65	65	0	72.5	58	14.5	88	90	2
	93	88	5	88	88	0	71	79	8	53	56	3	68	75	7	91	86	5	83	78	5
	89	85	4	78	76	2	96	87	9	75	75	0				62.5	60	2.5	93	93	0
	92	86	6	83	80	3	83	90	7	67	64	3				66	65	1	83	87	4
Aver. . .	86.8	84.2	4	85.5	84.7	3	70.3	74.3	7.8	71.9	72.2	2.1	76.0	78.4	2.8	77.1	71.1	6.5	82.0	85.2	4.6

The mean variation of these marks, comparing the first with the second for each paper, is 2.2 points. A part of this variation, however, is due to the slight shift in standard on the part of each teacher from one grading to the next. By applying the same process of weighting explained in connection with factor two, the mean variation drops from 2.2 to 1.75 points. Hence 1.75 represents the amount of variation contributed by factor four to the total mean variation and by subtraction we find that factor three contributes 2.55 points. The four factors therefore contribute the following amounts:

Factor one	.10	points
Factor two	1.00	"
Factor three	2.55	"
Factor four	1.75	"
Total mean variation	5.40	"

It is obvious then that factors three and four are the most important ones in producing the large differences of values assigned by teachers to a given piece of school work.

How Large Should the Units of a Marking Scale Be? The answer to this question depends primarily upon the fineness of the discriminations of successive degrees of quality in terms of the scale used, and, secondarily, upon the convenience of using a given scale of marks. The smallness of distinguishable shades of quality of anything can be determined by ascertaining the amount of difference in terms of a given scale that can be discriminated in the long run by the judges.

A general principle that has been followed commonly in psychological measurements, in which the units depend upon the discrimination of judges, has been to regard a difference in amount of the thing in question which can be distinguished correctly by 75% of the judges or judgments concerning it as the smallest psychological unit that can be used with reasonable certainty. For example, if we take ten shades of blue and ask 100 judges to arrange them in the order of blueness from left to right, we would regard that difference in blueness between any two successive shades which 75 of the judges agree in perceiving the one bluer than the other as the least that can be distinguished with a fair degree of confidence. The 75% point is chosen because it is midway between pure chance and absolute certainty. If only one of two possible judgments may be made, that is, if a given shade is either more or less blue than another, then 50% of the judgments would be correct by pure chance guesses. If 100% of the judgments are correct, it means that the difference is so large that it can be recognized correctly every time, and the amount of difference may range anywhere from just large enough to be always recognized up to an infinite difference.

According to this principle, how large would the steps be on the marking scale? For example, if we took two papers, *a* and *b*, in a given subject which differed in quality just enough so that three-fourths of the examiners or teachers would consider *b* better than *a*, how large would this difference be, say on the usual 100 percentage scale? Data for answering this question with approximate accuracy are found in Figures 87 to 90. The probable error or median deviation of the marks given by the teachers to the papers represented in these four figures are 4.0, 4.8, 7.8 and 7.1 respectively, with an average of 6.4. By definition, two times the probable error includes the middle half of the measures or marks. For example, in Figure 88 the median is 80.2 and the probable error is 4.8, that is, the middle 71 of the 142 marks lie between 75 and 85. Obviously one-fourth or 35 of the marks lie above 85. Consequently so far as this particular paper is concerned the next better paper would have to be 4.8 points better so that three-fourths of the examiners would consider it better.

Now the average probable error of the four sets of marks is 6.4. Hence the difference between two papers in general must be approximately 6.4 points so that three-fourths of the examiners would consider one better than another. On this principle then the step on the 100 percentage scale, with 70 as the usual passing grade, turns out to be approximately 7 points. This would produce a scale of steps as follows: 70–76, 77–84, 85–92, and 93–100. That is, the marking scale would have five steps, failure and four passing steps above 70. which may be designated as excellent, good, fair, poor, and failure, or perhaps preferably by the symbols A, B, C, D, and E. Such would be the size of the steps so that three-fourths of the examiners of a given set of papers would agree in distinguishing between the qualities of the papers.

However, any individual teacher agrees with himself more closely in re-grading a set of papers than he agrees with other teachers, as indicated in Table 137. This table shows that the probable error or median deviation of a given teacher's marks in re-grading his own papers is approximately 2 points. By the same reasoning the amount of difference in quality between two papers would have to be 2 points in order that an individual teacher would consider one paper better than another in three out of four independent markings. Hence the marking scale for an individual teacher, who grades papers from his own viewpoint and compares them only with his own judgments, could have each step in a five-step scale subdivided into three smaller steps of about 2 points each by using the plus and minus. That is 70 to 76 would become 70–72 or D –, 73–74 D, and 75–76 D+, and so on.

Whether a fine marking scale such as the 100 percentage scale or a coarser five-step scale should be used is largely a matter of convenience and personal habit. The advantage of a coarse scale is perhaps that it avoids giving the pupil the impression that the evaluation of a piece of work is more accurate than it actually is. The advantage of a fine scale is that it probably encourages the examiner in making as fine distinctions as possible. In practice a fine scale can probably be used as readily and as quickly as a coarse one if the teacher is accustomed to using a fine scale. A person may use as fine a scale as he wishes provided he recognizes the amount of the probable error in terms of the units of that particular scale. In terms of a 100 percentage scale the probable error is about 6 or 7 points; in terms of the five-step scale it is about one step which is 6 or 7 times as large as a point on the percentage scale. The absolute amount of variation is substantially the same on the two scales.

How Should Marks be Distributed to Groups of Pupils? If a five-step scale is used, what percentages of pupils should in the long run receive each of the five marks? The answer to this question that I advocate is that the marks of large numbers of unselected pupils should be distributed approximately in conformity with the normal distribution or probability curve. Three lines of evidence for this position may be presented, the last two of which are fundamentally based upon the first:

First, mental and physical traits, when measured in large numbers of individuals, are distributed in a manner which yields a distribution surface very nearly identical with that of the probability curve. Concrete evidence for this statement has been presented in Chapter III, Figures 7 to 10, to which the reader should turn. It seems reasonable to infer that abilities in school subjects are very probably distributed in the same manner as other mental traits.

In the second place, when abilities in school subjects are measured by objective methods, they are found to be distributed in very close conformity to the probability curve. Concrete evidence for this is presented in Figures 16 to 27, Chapter III.

Thus, for example, the scores of 662 seventh grade pupils in the author's geography test shows the following distribution when the total range of the base line is divided into five equal sections:

Scores............	0–27	28–54	55–81	82–108	109–135
% of pupils........	6	24	37	24	9

This is a remarkably close conformity to the theoretical distribution proposed on the following pages.

In the third place, the distribution of marks assigned by many teachers to large numbers of students conforms fairly closely to the normal distribution curve. When the marks of many teachers are combined, the idiosyncrasies of individual teachers tend to be counterbalanced. Tables 138 to 142, and Figures 91, 92, 93, and 94, show the distributions of marks in various institutions and the extent to which they differ from the theoretical probability curve.

TABLE 138

Distribution of grades in the College of Letters and Science, University of Wisconsin, for the years 1907, 1910 to 1915. From the reports of President E. A. Birge.

	Incomplete	Conditioned & Failed	Poor	Fair	Good	Excellent	No. of Grades
Elementary Course..	3.6	9.3	15.3	33.2	29.4	9.2	42,557
Advanced Courses...	3.2	3.5	7.9	30.9	41.8	12.7	39,302

TABLE 139

Distribution of grades at Cornell University for the years 1902, 1903 and 1911. Adapted from Finkelstein ('13, p. 22), to give the distributions for a five-point scale, 60 being the passing grade.

0–59	60–69	70–79	80–89	90–100	Number of Grades
9.2	22.5	30.0	27.2	11.1	20,348

TABLE 140

Distribution of all grades for two academic years at Harvard College. After Foster ('11, p. 262)

Totals	E%	D%	C%	B%	A%	Number of Grades
Elementary Courses..........	7	21	42	20	7	8969
Intermediate "	4	13	37	28	12	2456
Advanced "	2	2	13	38	36	476

TABLE 141

Distribution of grades at the University of Missouri. After Foster, p. 289

	Delayed	E	D	C	B	A	Number of Grades
Aug. 1908.....	3.5	15.6	8.7	41.2	23.3	7.7	
Feb. 1909.....	5.0	8.5	13.7	47.5	20.7	4.6	
June 1909.....	3.8	8.0	13.8	48.8	21.0	4.6	
Feb. 1910.....	3.5	6.5	14.4	49.6	21.3	4.7	24,979
Averages......	3.7	9.5	12.7	46.8	21.6	5.4	
First year after new system went into effect		9.0	14.5	50	21.7	4.9	11,342

TABLE 142

Average percentages for Cornell, Missouri, and the elementary courses for Harvard and Wisconsin. These percentages do not total 100 because the incomplete grades for Wisconsin and Missouri are not included.

E	D	C	B	A	Number of Grades
8.7	17.9	38.0	24.5	8.2	96,853

If we grant that marks in the long run should be assigned according to the normal distribution curve, what percentage of pupils should receive each of the five steps of the marking scale? If the base line of the probability curve in Figure 15, Chapter III, is divided into five equal divisions, then the area above the various divisions would comprise the following percentages of the total area:[1]

A, Excellent,	or 93–100 =	7%
B, Good,	or 85– 92 =	24%
C, Fair,	or 77– 84 =	38%
D, Poor,	or 70– 76 =	24%
E, Failure,	or 69 =	7%

Figures 91 to 94 indicate how closely distributions of the marks at Wisconsin, Cornell, Harvard, and Missouri run parallel to the theoretical curve. The only difference is a slight skewing to the right. Not quite as many D's are assigned and very slightly more E's and A's are assigned than the theoretical distribution would demand. Thus the marks as actually assigned by hundreds of

[1] The ends of the probability curve would reach the base line only at infinity. Hence an arbitrary point of termination must be selected. This has been placed at a point 3.65 P. E. values from the median. This point has been selected because it yields 7% for the E and A surfaces which is approximately the percentage of pupils receiving these grades in many institutions.

teachers to thousands of students furnish impressive support for the theory of the probability distribution of grades.

Certain objections, however, both of a theoretical and a practical kind, must be considered. In the first place, the soundness of

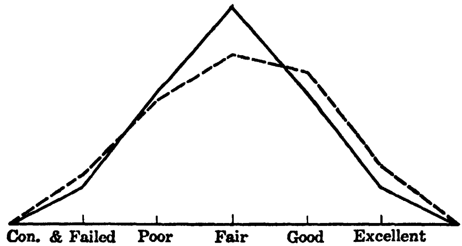

FIG. 91.—Distribution of 42,557 grades, broken line, in elementary courses in the College of Letters and Science of the University of Wisconsin. The continuous line is the theoretical distribution. After a report by President E. A. Birge.

the theory rests on the supposition that the pupils are unselected, chance specimens of mankind as a whole. This supposition, of course, never obtains absolutely for any group of human beings brought together anywhere. The very reason that brings any group together at the same time selects them. Pupils in school are

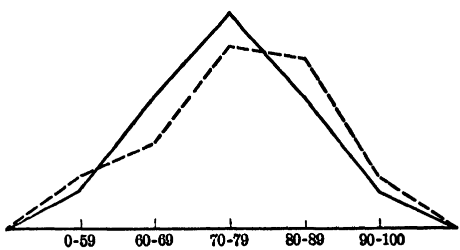

FIG. 92.—Distribution of 20,348 grades at Cornell University. After Finkelstein ('13).

not random samplings of human beings of their respective ages—the less so as one goes up the educational ladder. The tendency is that every rung of the ladder selects on the whole slightly better and better specimens. The fact, however, seems to be that the selection which does take place is not of the sort that materially modifies the form of the distribution curve but rather tends to contract its base. The selection that does take place is not an abrupt cutting off, but a gradual slicing off along a large share of the distribution surface.

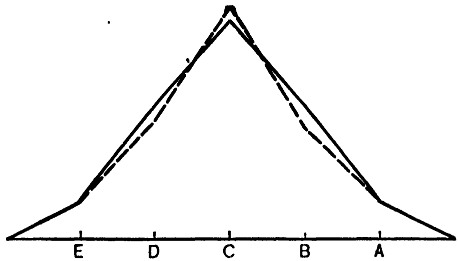

FIG. 93.—Distribution of 8,969 grades in elementary courses at Harvard University. After Foster ('11).

The writer undertook to ascertain the actual elimination of university students as it really takes place on the basis of the records of 476 freshmen tabulated by Dearborn. It was found that the

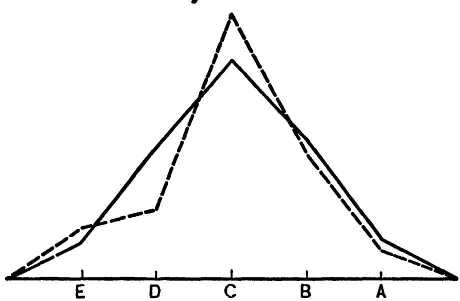

FIG. 94.—Distribution of 24,979 grades at the University of Missouri. After Meyer ('08).

following percentages of students dropped out of the University in the various grades of scholarship at the end of the freshmen and sophomore years:

TABLE 143

	CONDITIONED & FAILED	POOR	FAIR	GOOD	EXCELLENT
Percentage of students of each grade dropped during freshman year..................	100	52	19	11	0
Percentage of those remaining in each grade, dropped during sophomore year...........		45	16	8	0

This table reads that all students whose average grade was "conditioned" or "failed" dropped out during the freshman year; 52% of those whose grade was "poor" dropped out during the freshman year and 45% of those remaining whose grade was "poor" dropped out during the sophomore year, etc. It is obvious, therefore, that there is elimination from all classes of scholarship with the exception of the highest from which there is very little or no loss. The general effect of the actual elimination upon the distribution curve is to shift the left end of the curve toward the right and to change the general form of it only slightly as indicated in Figure 95.

The outcome of this evidence is that the distribution of the grades for the freshman year of the college as well as of the high

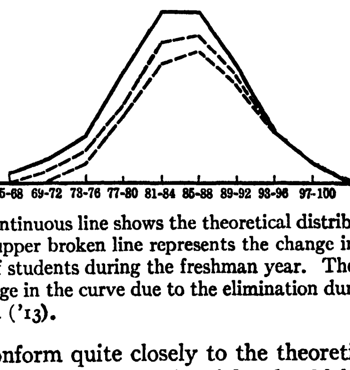

FIG. 95.—The continuous line shows the theoretical distribution of the marks of students. The upper broken line represents the change in this curve due to the dropping out of students during the freshman year. The lower broken line represents the change in the curve due to the elimination during the sophomore year. After Starch ('13).

school should conform quite closely to the theoretical distribution curve and that slight shifts to the right should be made for the successive four years. It may seem curious to recommend that after the elimination of the successive years of the high school, the distribution to be followed in the freshman year of the college should be approximately normal again. The explanation is that the standards of the college are somewhat higher than those of the high school; so that, even if the high school should eliminate all of the poorest 7% of its pupils, the next poorest 7%, who are able to complete the high school, are likely to be unable to meet the demands of the college.

The second objection urged by teachers against the adoption of the theoretical distribution of grades here recommended is that it would be unfair to lay down a rule that 7% of the pupils should be failed. How do we know; possibly by good teaching all pupils may reach a sufficiently high attainment to be passed in the course. The answer to this statement is that the effects even of the best teaching will so rarely raise the attainments of pupils sufficiently high so that none of the pupils would fall below the passing grade; and furthermore, in the interests of reasonably high standards of scholarship the attainments of approximately 7% of large numbers of pupils will very probably not merit a passing grade. There should be doubly good reasons for passing all students or for failing considerably less than 7%. Many of the cases of "good teaching" or "unusual classes" prove to be spurious when it is possible to check them up by outside means.

A third point is not so much an objection as a question of practical use of the principle of the distribution of grades; namely, in how large classes or groups of pupils should we expect fairly close conformity, and how close conformity should be expected? The answer to the question which I shall give, on the basis of experience in attempting to observe the principle in the assignment of grades, is that for groups of students of 100 or more quite close conformity should be expected. By quite close conformity I mean a deviation of not more than about 25% above or below the number of grades that theoretically should fall on a given step of the five-division scale. For example, the theoretical distribution demands that 7% of the pupils should receive the grade of A or Excellent. For groups of 100 or more pupils, this percentage should ordinarily not be lower than 5 nor higher than 9; the percentage of B's should ordinarily not run lower than 18 nor higher than 30; the percentage of C's should not run higher than 48, nor lower than 28, etc. The larger the number of pupils concerned, the closer the conformity should be. For groups smaller than 100 a wider latitude should be permissible whenever there is genuine reason for wider deviation. I advocate conformity to the theoretical distribution within the limits of common sense with as much deviation as may seem permissible for good cause. However, really genuine reasons for large deviations, even with classes as small as 25 pupils, unless obviously selected by special cause, is much rarer than teachers ordinarily believe.

In support of this contention, the author ('15) reported an experiment in which twenty-four compositions written by sixth and seventh grade pupils were graded by 23 teachers according to the usual percentage method with 70 as a passing grade. After the papers had thus been graded, the teachers were requested to grade them according to a five point scale and give the grade of E to two papers, D to from four to six papers, C to from eight to ten papers, B to from four to six papers, and A to two papers. Even if the teachers felt, for example, that there were no papers good enough to receive the grade of A, they were to select the two best ones and call them A. The outcome was that those teachers who in their original grades differed most from the combined judgment of all the teachers were forced to comply more closely to the actual average marks as given in the first grading. One teacher marked the highest paper 85 in the original grading, and objected to giving it a grade of A in the forced distribution on the ground that no paper in the lot was good enough to receive so high a grade, and yet the average of the marks given by all the teachers to this paper was 92.9, the best paper in the entire group.

The theory of the probability distribution of marks should be observed with sense and reason and not in a purely mechanical manner. A blind, unintelligent observance of the principle is bound to lead to injustice, particularly with small classes. In one such case which came to the author's attention it led to the giving of a mediocre grade to a pupil of very high ability.

A fourth point frequently raised by teachers to justify unusually high or low marks is that the particular class in question is an unusually good one or poor one. Such a claim ought to be allowed only if it can be justified by good evidence. There are, of course, differences in classes, but these are almost never as great as we are inclined to believe. Large differences between successive classes in the same subject are for the most part illusory for the reason that the judgment of an individual teacher is more likely to deviate from a correct estimate than the average ability of a group deviates from the average of other groups. The teacher who says to each succeeding class that this is the best class he has ever had in this subject would possess, if this judgment were correct, a magic power for elevating the intellectual level of human beings.

The feeling on the part of teachers that a given class is an unusually good or poor class is quite often due to one or two unusually good or poor individuals whose impression upon the mind of the teacher is outstanding, rather than to a higher or lower level of the class as a whole.

As concrete evidence of the extent to which a teacher may err in such opinions and of the manner in which the opinion of the teacher may be checked up, the curves in Figure 96 are presented.

The continuous curve shows the distribution of the grades of a teacher in Latin and German. When her attention was called to the predominance of high marks, she claimed that her pupils were exceptionally good. The broken curve shows the extent to which her claim was unfounded since it shows that, according to their abilities in other subjects, they were an average group.

How May Variation in the Assignment of Grades be Reduced.

1. By a common sense compliance in the distribution of marks with the normal distribution or probability curve.

(a) To this end the administrator of a school should tabulate at stated intervals the marks assigned by each teacher and exhibit the tabulation to the teachers. This in itself will usually lead without request or compulsion to a very considerable correction

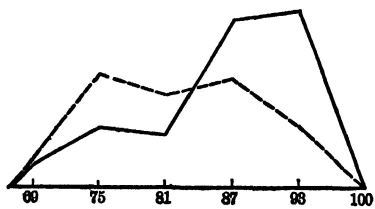

FIG. 96.—The continuous line shows the distribution of the marks of a teacher of Latin and German in a high school. The broken line shows the distribution of the marks of the same pupils in their other subjects. After an unpublished report of Supt. J. F. Waddell, Evansville, Wisconsin.

of aberrations on the part of those teachers who deviate most widely. At the University of Missouri the adoption of a plan of distribution in conformity with the normal curve reduced the irregularity in grading in the ratio of five to two.

(b) The teacher himself will find it useful to tabulate at frequent intervals the distribution of his grades. In making out the marks of a set of papers and particularly in making out the final grades for a course, the author has followed for several years the practice of plotting a distribution of the grades as tentatively made out. If the assignment is decidedly abnormal in having considerably too many or too few of the different grades, a shift is made of borderline cases, unless there is an obvious reason to the contrary, to obtain a reasonably normal distribution. Every teacher feels that there is a considerable number of cases concerning which he is in doubt as to whether they should have the one or the other grade. For example, if the tentative list of grades contains too many or too few A's, the lowest A's may be shifted to B's or the highest B's may be shifted to A's.

2. Variability and uncertainty in grades may be reduced by adopting particularly in departments containing several teachers, a plan of giving certain weights or penalties for certain types of errors or defects. This should be done by departmental conference so as to secure a consensus of judgments on the various types of errors and amounts of penalties. Much could be done in this direction toward greater uniformity in methods of grading. If organizations of teachers would take this matter up, much of the chaos which is now present in methods of grading could be reduced to order.

At the present time A's or B's obtained from different teachers often mean quite different things. By observing the points here suggested they would mean more nearly the same thing. Evaluation of achievement in terms of judgment depends obviously upon the judge. Marks as such will at best depend upon the examiner. They will probably always have to be used. More impersonal and objective methods for determining achievement in school work are being developed at the present time. To what extent these educational measuring devices will be able to replace the usual examinations and grades will depend upon their future development.

BIBLIOGRAPHY

Abbott, E. E.: 1909. On the Analysis of the Memory Consciousness in Orthography. Psychological Review Monograph, 11: No. 1, 127–158.

Arai, T.: 1912. Mental Fatigue. Columbia University Contributions to Education, No. 54.

Ash, I. E.: 1914. Fatigue and Its Effects upon Control. Archives of Psychology, No. 31.

Ayres, L. P.: 1912. A Scale for Measuring the Quality of Handwriting of School Children. Russell Sage Foundation, New York City.

Ayres, L. P.: 1913. The Spelling Vocabularies of Personal and Business Letters. Russell Sage Foundation, New York City.

Ayres, L. P.: 1915. A Measuring Scale for Ability in Spelling. Russell Sage Foundation, New York City.

Bagley, W. C.: 1911. Educational Values.

Bagley, W. C.: 1915. The Determination of Minimum Essentials in Elementary Geography and History. Fourteenth Yearbook of the National Society for the Study of Education, 131–146.

Bagley, W. C., and Rugg, H. O.: 1916. The Content of American History as taught in the Seventh and Eighth Grades: An Analysis of Typical School Text-books. University of Illinois Bulletin.

Bair, J. H.: 1902. The Practice Curve. Psychological Review Monograph, No. 19.

Baldwin, B. T.: 1914. United States Bureau of Education, Bulletin No. 581.

Ballard, P. B.: 1915. Norms of Performance in Fundamental Processes of Arithmetic, Journal of Pedagogy, 3:9–20.

Ballou, F. W.: 1914. Scales for the Measurement of English Composition. Harvard University.

Batson, W. H.: 1916. Acquisition of Skill. Psychological Review Monographs, 21, No. 3.

Bean, C. H.: 1912. The Curve of Forgetting. Archives of Psychology, No. 21.

Bennett, C. J. C.: 1907. Formal Discipline. Columbia University.

Bennett, N. F.: 1916. The Correlations Between Different Memories. Journal of Experimental Psychology, 1:404–418.

Bergstrom, J. A.: 1894. The Relation of the Interference to the Practice Effect of an Association. American Journal of Psychology, 6:433–442.

Betts, G. H.: 1909. The Distribution and Functions of Mental Imagery. Columbia University Contributions to Education, No. 26.

Betts, G. H.: 1917. Class-Room Method and Management.

Binet, H., and Simon, T.: 1905. (a) Sur la nécessité d'établér un diagnostic scientific des états inferieurs de l'intelligence. Année Psychologique, 11:163–190. (b) Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. Ibid., 191–244. (c) Application des méthodes nouvelles an diagnostic du niveau intellectuel chez enfants normaux et anormaux d'hospice et d'école primaire. Ibid., 245–336.

Boas, F.: 1896–1897. Growth of American Children. Report of United States Commissioner of Education, 11:1555.

Bolton, T. L.: 1902. Ueber die Beziehungen zwischen Ermüdung, Roumsinn der Haut and Muskelleistung. Psychologische Arbeiten, 4:175–234.

Book, W. F.: 1908. 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.

Borst, M.: 1904. Recherches exp'tlles sur l'educabilité et la fidelite du temoignoge. Archives de Psychologie, 3:233–314.

Breslich, E. R.: 1912. Teaching High School Pupils to Study. School Review, 20:505–515.

Briggs, T. H.: 1913. Formal English Grammar as a Discipline. Teachers' College Record, 14, No. 41.

Briggs, T. H.: '13. The Right Way to Read. Education, 33:559–562.

Brown, H. A.: 1914. Measurement of Efficiency of Instruction in Reading. Elementary School Teacher, 14:477–490.

Brown, H. A.: 1916. The Measurement of Ability to Read. Department of Public Instruction, New Hampshire, Bulletin No. 1.

Brown, J. C.: 1913. An Investigation on the Value of Drill Work in the Fundamental Operations in Arithmetic. Journal of Educ. Psych., 2:81–88; 3:561–570.

Brown, M. D., and Haggerty, M. E.: 1917. The Measurement of Improvement in English Composition. The English Journal, 6:515–527.

Browne, C. E.: 1906. The Psychology of the Simple Arithmetical Processes. American Journal of Psychology, 17:1–37.

Bryan, W. L., and Harter, N.: 1897–1899. Studies in the Physiology and Psychology of the Telegraphic Language. Psychological Review, 4:27–53; 6:348–375.

Buckingham, B. R.: 1913. Spelling Ability: Its Measurement and Distribution. Teachers' College Record.

Burk, C. F.: 1900. The Collecting Instinct. Pedagogical Seminary, 7:179–207.

Burnett, C. J.: 1906. The Estimation of Number. Harvard Psychological Studies, 2:349–404.

Burt, C.: 1909. Experimental Tests of General Intelligence. British Journal of Psychology, 3:94–177.

Carter, M. H.: 1909. The Conservation of the Defective Child, McClure's Magazine, 33:160.

Cattell, J. M.: 1886. Inertia of the Eye and Brain. Brain 8:294–309.

Cattell, J. M.: 1886. Ueber die Zeit des Erkennens, u. s. w. Philos. Studien, 3:94–127.

Cattell, J. M.: 1905. Examinations, Grades and Credits. Popular Science Monthly, 66:367–378.

Cattell, J. M.: 1906. A Statistical Study of American Men of Science, III. The Distribution of American Men of Science. Science, New Series, 24:723–742.

Charters, W. W., and Miller, Edith: 1915. A Course of Study in Grammar. University of Missouri, Education Bulletin, No. 9.

Colvin, S. S., and Myers, E. J.: 1909. The Development of Imagination in School Children and the Relation between Ideational Types and the Retentivity of Material Appealing to Various Sense Departments. Psychological Review Monographs, 11: No. 1, 85–126.

Conrad, H. E., and Arps, G. F.: 1916. An Experimental Study in Economical Learning. American Journal of Psychology, 27:507–529.

Cook, W. A.: 1912. Shall we Teach Spelling by Rule? Journal of Educational Psychology, 3:316:325.

Cook, W. A., and O'Shea, M. V.: 1914. The Child and His Spelling.

Coover, J. E.: 1916. Formal Discipline from the Standpoint of Experimental Psychology. Psychological Review Monographs, 20: No. 3.

Coover, J. E., and Angell, F.: 1907. General Practice Effect of Special Exercise. American Journal of Psychology, 18:327–340.

Cornell, W. S.: 1908. The Relation of Physical to Mental Defect in School Children. Psychological Clinic, 1:233.

Cornman, O. P.: 1902. Spelling in the Elementary School.

Counts, G. S.: 1917. Arithmetic Tests and Studies in the Psychology of Arithmetic. Suppl. Educ. Monographs, 1: No. 4. University of Chicago.

Courtis, S. A.: 1910. Manual of Instruction. 82 Eliot Street, Detroit, Michigan.

Dallam, M. T.: 1917. Is the Study of Latin Advantageous to the Study of English? Educational Review, 54:500–503.

Davidson, P. E.: 1914. The Recapitulation Theory and Human Infancy. (Columbia University.)

Davis, W. W.: 1898–1900. Researches in Cross-Education. Studies from the Yale Psychological Laboratory, 6:5, 6, and 8, 54–108.

Dearborn, W. F.: 1906. The Psychology of Reading. Archives of Psychology, No. 4.

Dearborn, W. F.: 1909. The General Effects of Special Practice in Memory. Psychological Bulletin, 6:44.

Dearborn, W. F.: 1909. The Relative Standing of Pupils in the High School and the University. Bulletin of the University of Wisconsin, No. 312.

Dearborn, W. F.: 1910. Experiments in Learning. Journal of Educational Psychology, 1:373–388.

Dearborn, W. F.: 1910. School and University Grades. Bulletin of the University of Wisconsin, No. 368.

De Candolle, A.: 1873. (Second Edition, 1885.) Histoire des sciences et des savants depuis deux siecles, etc.

Delabarre, E. B.: 1898. A Method of Recording Eye Movements. American Journal of Psychology, 9.

Dietze, G.: 1885. Ueber den Bewusstseinsumfang.

Dodge, R.: 1900. Visual Perception During Eye-movements. Psychological Review, 7:454–465.

Dougherty, M. L.: 1913. Report of the Binet-Simon Tests Given to 483 Children in the Public Schools of Kansas City, Kansas. Journal of Educational Psychology, 4:338–352.

Downey, J. E.: 1908. Control Processes in Modified Handwriting: An Experimental Study. Psychological Review Monographs, 9, No. 1.

Downey, J. E.: 1910. Judgments on the Sex of Handwriting. Psychological Review, 17:205–216.

Drushel, J. A.: 191–. A Study of the Amount of Arithmetic at the Command of High School Graduates who have had no Arithmetic in their High School Course. El. School Journal, 17:657–661.

Earle, E. L.: 1903. The Inheritance of the Ability to Learn to Spell. Columbia Contrib. to Phil., Psy., and Ed. 2, No. 2:41–44.

Ebbinghaus, H.: 1885. Ueber das Gedächtniss.

Ebbinghaus, H.: 1897. Ueber eine neue Methode zur Prüfung geistiger Fähigkeiten und ihre anwendung bei Schulkindern. Zeitschrift für Psychologie, 13:401–459.

Ebert, E., and Meumann, E.: 1905. Ueber einige Grundfragen der Psychologie des Gedächtnisses. Archiv für gesamte Psychologie, 4:1–32.

Eldridge, R. C.: 1911. Six Thousand Common English Words. Niagara Falls, New York.

Erdmann, B., and Dodge, R.: 1908. Psychologische Untersuchungen über das Lesen, auf Experimenteller Grundlage. Halle.

Fillers, H. D.: 1917. Oral and Written Errors in Grammar. Educational Review, 54:458–470.

Finkelstein, I. E.: 1913. The Marking System in Theory and Practice.

Foster, F. M.: 1917. The Results of a Recent Spelling Test at the University of Iowa. School and Society, 5:506–508.

Foster, W. T.: 1911. Administration of the College Curriculum.

Fracker, G. C.: 1908. On the Transference of Training in Memory. Psychological Review Monographs, No. 38:56–102.

Freeman, F. N.: 190–. Untersuchungen über den Aufmerksamkeitsumfang und die Zahlanffassung. Pädagogische-Psychologische Arbeiten, 1:88–168.

Freeman, F. N.: 1914. Practical Studies of Handwriting. Elementary School Teacher, 14:167–179.

Freeman, F. N.: 1914. The Teaching of Handwriting.

Freeman, F. N.: 1915. Handwriting. Fourteenth Yearbook of the National Society for the Study of Education.

Freeman, F. N.: 1916. The Psychology of the Common Branches.

Friedrich, J.: 1897. Untersuchungen über die Einflüsse der Arbeitsdauer und der Arbeitspausen auf die geistige Leistungsfähigkeit der Schulkinder. Zeitschrift für Psychologie, 13:1–53.

Fulton, M. J.: 1914. An Experiment in Teaching Spelling. Ped. Sem., 21:287–289.

Galton, F.: 1869 and 1892. Hereditary Genius: An Inquiry into its Laws and Consequences.

Galton, F.: 1883. Inquiries into Human Faculty and its Development.

Gesell, A. L.: 1906. Accuracy in Handwriting, as Related to School Intelligence and Sex. American Journal of Psychology, 17:394–405.

Gilbert, J. A.: 1894. Researches on the Mental and Physical Development of School Children. Studies from the Yale Psychological Laboratory, II:40–100.

Gilbert, J. A., and Fracker, G. C.: 1897. The Effects of Practice in Reaction and Discrimination for Sound upon the Time of Reaction and Discrimination for other Forms of Stimuli. University of Iowa Studies in Psychology, 1:62–76.

Gist, A. S.: 1917. Errors in the Fundamentals of Arithmetic. School and Society, 7:175–177.

Goddard, H. H.: 1912. The Kallikak Family.

Goddard, H. H.: 1914. Feeble-mindedness: Its Causes and Consequences.

Graves, S. M.: 1917. A Study in Handwriting. J. of Educ. Psych., 7:483–494.

Gray, C. T.: 1913. Variations in the Grades of High School Pupils.

Gray, C. T.: 1915. A Score Card for the Measurement of Handwriting. Bulletin of the University of Texas, No. 37.

Gray, W. S.: 1917. Studies of Elementary School Reading Through Standardized Tests. Suppl. Educ. Monog., 1, No. 1.

Gulick, L. H., and Ayres, L. P.: 1908. (Second Edition, 1913.) Medical Inspection of Schools.

Guyer, M. F.: 1916. Being Well-Born.

Haggerty, M. E.: 1917. Scales for Reading Vocabulary of Primary Children. Elementary School Journal, 17:104–115.

Hahn, W. H., and Thorndike, E. L.: 1914. Some Results in Addition under School Conditions. Journal of Educ. Psych., 5:65–84.

Haines, T. H.: 1916. Mental Measurements of the Blind. Psych. Rev. Monog., 21, No. 1.

Hall, G. S.: 1883. Published first in Princeton Review, 2:249–272. Then reprinted in "Aspect of Child Life and Education," 1907.

Hall, G. S.: 1904. Adolescence.

Hall, G. S.: 1908. Youth.

Hall-Quest, A. L.: 1917. Supervised Study.

Hamilton, F. M.: 1907. The Perceptual Factors in Reading. Archives of Psychology, No. 9.

Harris, J. H. and Anderson, H. W.: 1916. Measuring Primary Reading in Dubuque Schools (Iowa).

Harris, L. H.: 1915. A Study in the Relation of Latin to English Composition. School and Society, 2:251–252.

Heck, W. H.: 1913. A Study of Mental Fatigue.

Heck, W. H.: 1913. A Second Study of Mental Fatigue in Relation to the Daily School Program. Psychological Clinic, 7:29–34.

Henmon, V. A. C.: 1910. Sex Differences and Variability in Color Perception. University of Colorado Studies, 7, No. 4.

Henmon, V. A. C.: 1912. The Relation between Mode of Presentation and Retention. Psychological Review, 19:79–96.

Hewins, N. P.: 1916. The Doctrine of Formal Discipline in the Light of Experimental Investigation.

Hillegas, M. B.: 1912. A Scale for the Measurement of Quality in English Composition by Young People. Teachers' College Record, 13, No. 4.

Hoge, Mildred, and Stocking, Ruth, J.: 1912. A Note on the Relative Value of Punishment and Reward as Motives. Journal of Animal Behavior, II, 43.

Hollingworth, H. L.: 1912. Correlation of Abilities as Affected by Practice; Journal of Educational Psychology, Jan., 1912.

Hollingworth, H. L.: 1914. Individual Differences, During and After Practice. Psychological Review, January, 1914.

Holloway, H. V.: 1917. The Relative Difficulty of the Tables, Trenton, N. J.

Horn, E.: 1917. Possible Defects in the Present Content of American History as Taught in the Schools. Sixteenth Yearbook of the National Society for the Study of Education.

Hosic, J. F.: 1915. The Essentials of Composition and Grammar. Fourteenth Yearbook of the National Society for the Study of Education.

Houser, J. D.: 1917. An Investigation of the Writing Vocabularies of Representatives of an Economic Class. Elem. School Journal, 17:708–718.

Howell, H. B.: 1914. A Fundamental Study in the Pedagogy of Arithmetic.
Hoyt, F. S.: 1906. Studies in the Teaching of English Grammar. Teachers' College Record, 7, No. 5.
Huey, E. B.: 1898 and 1900. Experiments in the Physiology and Psychology of Reading. American Journal of Psychology, 9:575–586; 11:283–302; 12:292–312.
Huey, E. B.: 1910. The Psychology and Pedagogy of Reading.
Hull, C. L.: 1919. Quantitative Aspects of the Evolution of Concepts. Psychological Review Monograph Supplements, 1919.
Hull, C. L., and Montgomery, R. B.: 1919. An Experimental Investigation of Certain Alleged Relations between Character and Handwriting. Psych. Rev., 26:63–74.

Jacoby, P.: 1881. (Second Edition, 1904.) Études sur la selection dous ses rapports avec l'hérédité chez l'homme.
James, W.: 1890. Principles of Psychology.
Javal, E.: 1878 and 1879. Sur la Physiologie de la Lecture. Annoles d' Oculistique.
Javal, E.: 1879. Conditions de la lecture facile. Comptes rendus de la Societie de Biologie, page 8.
Jessup, W. A.: 1915. Current Practices and Standards in Arithmetic. Fourteenth Yearbook of the National Society for the Study of Education, 14:116–130.
Johnson, R. I.: 1917. The Persistency of Errors in English Composition. School Review, 25:555–580.
Jones, W. F.: 1913. Concrete Examination of the Material of English Spelling.
Jones, W. F.: 1915. Spelling Book.
Jost, A.: 1897. Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen. Zeitschrift für Psychologie, 14: 436–472.
Judd, C. H., and McAllister, C. N., and Steele, W. M.: 1905. Analysis of Reaction Movements. Psychological Review Monographs, 7:141–184.
Judd, C. H.: 1908. The Relation of Special Training to General Intelligence. Educational Review, 36:28–42.
Judd, C. H.: 1915. The Psychology of High School Subjects.
Judd, C. H.: 1916. Measuring the Work of the Public Schools.
Judd, C. H. (Edited by): 1916. School Survey, Grand Rapids, Michigan.

Kelley, T. L.: 1914. Educational Guidance. Columbia University.
Kelly, F. J.: 1915. The Kansas Silent Reading Test. Bulletin of State Normal School, Emporia, Kansas.
Kerr, Mary A.: 1916. The Effects of Six Weeks of Daily Drill in Addition. Indiana University Studies.

King, I.: 1917. Comparison of Rapid and Slow Readers in Comprehension. School and Society, 4:330–334.

Kirby, T. J.: 1913. The Results of Practice under School Conditions. Columbia University.

Kirkpatrick, E. A.: 1914. An Experiment in Memorizing vs. Incidental Learning. Journal of Educ. Psych., 5:405–413.

Kline, L. W.: 1909. Some Experimental Evidence on the Doctrine of Formal Discipline. Bulletin of State Normal School, Duluth, Minn.

Kraepelin, E., and Hoch, A.: 1895. Ueber die Wirkung der Theebestandtheile auf körperliche und geistige Arbeit. Psychologische Arbeiten, 1:378–488.

Kraepelin, E., and Hylan, J. P.: 1902. Ueber die Wirkung kurzer Arbeitszeiten. Psychologische Arbeiten, 4:454–494.

Kraepelin, E., and Rivers, W. H. R.: 1896. Ueber Ermüdung und Erholung. Psychologische Arbeiten, 1:378–488.

Kraepelin, E.: 1902. Die Arbeitscurve. Philsophische Studien, 19:459–507.

Kraepelin, E.: 1903. Ueber Ermüdungsmessungen. Archive für die gesamte Psychologie, 1:9–30.

Kuhlmann, F.: 1912. Revision of the Binet-Simon System for Measuring the Intelligence of Children. Journal of Psycho-Asthenics, Monograph Supplements, 1, No. 1.

Laser, H.: 1894. Ueber geistige Ermüdung beim Schulunterricht. Zeitschrift für Schulgesundheitspflege, 7:2–22.

Lay, W. A.: 1898. Führer durch den ersten Rechenunterricht. Wiesbaden.

Lay, W. A.: 1907. Führer durch den Rechenunterricht der Unterstufe.

Laudolt: 1891. Nouvelles recherches sur la physiologie des mouvements des yeux. Archives d'Ophthalmologie, 2:385–395.

Lindley, E. H.: 1897. A Study of Puzzles. American Journal of Psychology, 8:431–493.

Llewelyn, E. J.: 1916. Reading in the Mt. Vernon Schools. Elementary School Journal, 17:123–127.

Lodge, G.: 1915. Chapter X. Principles of Secondary Education. Edited by Paul Monroe.

Loisette, A.: 1899. Assimilative Memory.

Lowell, A. L.: 1910. College Studies and the Professional School. Harvard Graduates' Magazine, Dec., 1910; also Educational Review, Oct., 1911.

Lueba, J. H., and Hyde, W.: 1905. An Experiment in Learning to Make Hand Movements. Psychological Review, 12:351–369.

Lyon, D. O.: 1913. Relation of Length of Material to Time Taken for Learning and the Optimum Distribution of Time. Journal of Educational Psychology, 5:155–163.

Matthews, B.: The Study of Fiction in Counsel upon the Reading of Books by Henry VanDyke.

Mead, C. D.: 1915. Silent versus Oral Reading with 100 Sixth Grade Children. Journal of Educational Psychology, 6:345–348.

Mead, C. D.: 1917. Results of Silent versus Oral Reading. Journal of Educational Psychology, 8:367–368.

Mead, C. D., and Sears, Isabel. 1916: Additive, Subtraction and Multiplicative Division Tested. Journal of Educ. Psych., 7:261–270.

Messenger, J. F.: 1903. The Perception of Number. Psychological Review Monographs, 5:123–144.

Meumann, E.: 1914. (First Edition, 1907.) Vorlesungen zur Einfürhung in die Experimentelle Pädagogik und ihre Psychologischen Grundlagen. Vols. I, II, and III.

Meyer, M.: 1908. The Grading of Students. Science, 28:243–250.

Minnick, J. H.: 1913. An Experiment in the Supervised Study of Mathematics. School Review, 21:670–675.

Monroe, W. S.: 1917. A Preliminary Report of an Investigation of the Economy of Time in Arithmetic. Sixteenth Yearbook of the National Society for the Study of Education, 1917:111–127.

Mosso, A.: Fatigue.

Munn, A. F.: 1909. The Curve of Learning. In Studies in Development and Learning. Edited by E. A. Kirkpatrick, Archives of Psychology, No. 12.

Nanu, H. A.: 1904. Zur Psychologie der Zahl-Auffassung. Wurzberg.

Nicholson, F. W.: 1915. Success in College and in After Life. School and Society, 2:229–232.

Oberholtzer, E. E.: 1914. Testing the Efficiency of Reading in the Grades. Elementary School Journal, 15:313–322.

Partridge, E. A.: 1915. High-School Latin—A New Phase of an Old Subject. Classical Journal, 10:404–412.

Pearson, K.: 1904. On the Laws of Inheritance in Man. II. On the Inheritance of the Mental and Moral Characters in Man, etc. Biometrika, 3:131–190.

Pearson, K.: 1907. On the Relationship of Intelligence to Size and Shape of Head, and to Other Physical and Mental Characters. Biometrika, 5:105–146.

Perkins, A. S.: 1914. Latin as a Vocational Study in the Commercial Course. Classical Journal, 10:7–16.

Peters, C. C.: 1917. Influence of Speed Drills on Silent Reading. Journal of Educational Psychology, 8:350–366.

Peterson, H. A.: 1912. Note on a Retrial of Professor James' Experiment on Memory Training. Psych. Rev., 19:491–492.

Phelps, C. L.: 1913. A Study of Errors in Tests of Adding Ability. Elem. School Journal, 14:29–39.

Phelps, W. L.: Teaching in School and College.

Phillips, F. M.: 1913. Value of Daily Drill in Arithmetic. Journal of Educ. Psych., 4:159.

Pintner, R.: 1913. Oral and Silent Reading of Fourth Grade Pupils. Journal of Educational Psychology, 4:333–337.

Pintner, R., and Paterson, D. G.: 1916. Psych. Rev. Monog., 20: No. 4.

Pintner, R., and Paterson, D. G.: 1917. A Scale of Performance Tests.

Proctor, M.: 1917. Supervised Study on the Pacific Coast. School and Society, 6:326–327.

Pryor, H. C.: 1915. Spelling. Fourteenth Yearbook of the National Society for the Study of Education, 1915:78–89.

Puffer: 1912. The Boy and His Gang.

Pyle, W. H.: 1913. Economical Learning. Psychological Bulletin, 10:73.

Radassawljewitsch, P. R.: 1907. Das Behalten und Vergessen bei Kindern und Erwachsenen nach experimentellen Untersuchungen.

Rice, J. M.: 1897. The Futility of the Spelling Grind. The Forum, 23: 163–172, 409–419.

Rice, J. M.: 1902. Educational Research: A Test in Arithmetic. The Forum, 34:281–297; Causes of Success and Failure in Arithmetic. The Forum, 34:437–452.

Riley, J. L.: 1908. The Springfield Tests, 1846–1905, 1906.

Ritter, C.: 1900. Ermüdungsmessungen. Zeitschrift für Psychologie, etc., 24:401–444.

Ruediger, 1907. The Field of Distinct Vision. Archives of Psychology, No. 5.

Ruger, H. A.: 1910. The Psychology of Efficiency. Archives of Psychology, No. 15.

Rugg, H. O.: 1916. The Experimental Determination of Mental Discipline in School Studies.

Sackett, L. W.: Measuring a School System by the Buckingham Spelling Scale. School and Society, 2:860–864, 894–898.

Schmidt, W. A.: 1917. An Experimental Study in the Psychology of Reading. Suppl. Educational Monographs (Chicago), 1, No. 2.

Schuster and Elderton: 1907. Inheritance of Ability. Publications of Eugenics Laboratory, London.

Schuyler, W., and Swift, E. J.: 1907. The Learning Process. Psychological Bulletin, 4:307–310.

Scripture, E. W., Smith, T. L., and Brown, E. M.: 1894. On the Education of Muscular Control and Power. Studies from the Yale Psychological Laboratory, 2:114–119.

Sears, J. B.: 1915. Spelling Efficiency in the Oakland Schools. School and Society, 2:531–537, 569–574.

Sears, Isabel, and Diebel, Amelia: 1916. A Study of Common Mistakes in Pupils' Oral English. El. School Journal, 17:44-54.

Seashore, C. E.: 1901. Suggestions for Tests on School Children. Educational Review, 22:69-82.

Seashore, C. E., and Kent, G. H.: 1905. Periodicity and Progressive Change in Continuous Mental Work. Psychological Review Monographs.

Sikorski, J.: 1879. Sur les effects de la lassitude provoquee par les travoux intellectuels chez enfants le l'age scollaire. Annales d'hygiene publique, 2:458-464.

Simpson, B. R.: 1912. Correlations of Mental Abilities. Columbia University.

Sleight, W. G.: 1911. Memory and Formal Training. British Journal of Psychology, 4:386-457.

Smledey, F.: 1900-1901. Rept. Dept. Child Study and Pedagogic Investigation, No. 3.

Smith, F. O: 1912. A Rational Basis for Determining Fitness for College Entrance. University of Iowa Studies in Education, N. S. 51.

Sneed, C. M., and Whipple, G. M.: 1909. An Examination of the Eyes, Ears, and Throats of Children in the Public Schools of Jefferson City, Missouri. Psychological Clinic, 2:234-238.

Spillman, W. J.: 1909. The Country Boy. Science, 29:739-741; 30:405-407.

Starch, D.: 1910. A Demonstration of the Trial and Error Method of Learning. Psychological Bulletin, 7:20-23.

Starch, D.: 1911. Experiments in Educational Psychology, 1911. Revised, 1917.

Starch, D.: 1911. Unconscious Imitation in Handwriting. Psychological Review, 18:223-228.

Starch, D.: 1911. The Transfer of Training in Arithmetical Operations. Journal of Educational Psychology, 2:306-310.

Starch, D.: 1912. Periods of Work in Learning. Journal of Educational Psychology, 3:209-213.

Starch, D.: 1913. The Measurement of Handwriting. Journal of Educational Psychology, 4:445-464.

Starch, D.: 1913. Reliability and Distribution of Grades. Science 38:630-636.

Starch, D.: 1913. Correlation Among Abilities in School Studies. Journal of Educational Psychology, 3:415-418.

Starch, D.: 1915. The Measurement of Efficiency in Reading. Journal of Educational Psychology, 6:1-24.

Starch, D.: 1915. The Measurement of Efficiency in Handwriting. Journal of Educational Psychology, 6:106-114.

Starch, D.: 1915. The Measurement of Efficiency in Spelling. Journal of Educational Psychology, 6:127-186.

Starch, D.: 1915. The Measurement of Ability in English Grammar. Journal of Educational Psychology, 6.

Starch, D.: 1915. Can the Variability of Marks be Reduced? School and Society, 2:242–243.

Starch, D.: 1915. The Inheritance of Abilities in School Studies. School and Society, 2:608–610.

Starch, D.: 1915. Some Experimental Data on the Value of Studying Foreign Languages. School Review, 23:697–703.

Starch, D.: 1916. The Measurement of Ability in Arithmetic. Journal of Educational Psychology, April, 1916.

Starch, D.: 1917. The Estimated Value of School Subjects. School and Society, 5:59–60.

Starch, D.: 1917. Further Experimental Data on the Value of Studying Foreign Languages. School Review, 25:243–248.

Starch, D.: 1917. Similarity between Brothers and Sisters in Mental Traits. Psychological Review, 24:235–238.

Starch, D., and Ash, I. E.: 1917. The Mental Work Curve. Psychological Review, 24:391–402.

Starch, D., and Elliott, E. C.: 1912. The Reliability of Grading High School Work in English. School Review, 20:442–457.

Starch, D., and Elliott, E. C.: 1913. The Reliability of Grading Work in Mathematics. School Review, 21:254–259.

Starch, D., and Elliott, E. C.: 1913. The Reliability of Grading Work in History. School Review, 21:676–681.

Stone, C. W.: 1908. Arithmetical Abilities and Some Factors Determining Them. Columbia Contributions to Education, No. 10.

Studebaker, J. W.: 1916. Economy Practice Exercises in Arithmetic. Scott, Foresman & Company.

Suzzallo, H.: 1910. The Teaching of Spelling.

Suzzallo, H., and Pearson, H. C.: 1913. Comparative Experimental Teaching in Spelling. Teachers' College Record, 13: No. 1.

Swift, E. J.: 1908. The Mind in the Making.

Swift, E. J.: 1912. Youth and the Race.

Taussig, A. E.: 1909. The Prevalence of Visual and Aural Defects among the Public School Children of St. Louis County, Missouri. The Psychological Clinic, 3:149–160.

Terman, L. M.: 1916. The Measurement of Intelligence.

Terman, L. M.: 1917. The Intelligence Quotient of Francis Galton in Childhood. American Journal of Psychology, 28:210.

Thompson (Woolley), H. B.. 1903. Psychological Norms in Men and Women. University of Chicago Contributions to Philosophy, IV, No. 1.

Thorndike, E. L.: 1900. Mental Fatigue, Psychological Review, 7:466–482, 547–579.

Thorndike, E. L.: 1905. Measurements of Twins. Archives of Phil. Psychology and Scientific Methods, No. 1.

Thorndike, E. L.: 1910. Handwriting. Teachers' College Record, 11: No. 2.

Thorndike, E. L.: 1910. Practice in the Case of Addition. American Journal of Psychology, 21:483–486.

Thorndike, E. L.: 1911. Mental Fatigue. Journal of Educational Psychology, 12:61–80.

Thorndike, E. L.: 1912. Education.

Thorndike, E. L.: 1912. The Permanence of Interests and Their Relation to Abilities, Popular Science Monthly, Nov., 1912.

Thorndike, E. L.: 1912. The Curve of Work. Psychological Review, 19:165–194.

Thorndike, E. L.: 1914. Educational Psychology, Vol. I. The Original Nature of Man. Vol. II. The Psychology of Learning. Vol. III. Mental Work and Fatigue and Individual Differences and Their Causes.

Thorndike, E. L.: 1914. The Measurement of Ability in Reading. Teachers' College Record, 15, No. 4.

Thorndike, E. L.: 1915. The Disciplinary Values of Studies: A Census of Opinions. Education, 35:278–286.

Thorndike, E. L.: 1915. The Relation between Speed and Accuracy in Addition. Jour. of Educ. Psych., 5:537.

Thorndike, E. L.: 1917. The Curve of Work and the Curve of Satisfyingness, Journal of Applied Psychology, 1:265–267.

Trabue, M. R.: 1916. Completion-test Language Scales. Columbia University Contributions to Education, No. 77.

Uhl, W. T.: 1916. Use of the Results of Reading Tests as Basis for Planning Remedial Work. Elementary School Journal, 17:266–275.

Waldo, K. D.: 1914. Tests in Reading in the Sycamore Schools, Elementary School Journal, 15:251–268.

Wallin, J. E. W.: 1911. Spelling Efficiency in Relation to Age, Grade and Sex, and the Question of Transfer.

Walsemann, H. J.: 1907. Anschauungslehre der Rechenkunst.

Warren, H. C.: 1897. The Reaction Time of Counting. Psychological Review, 4:569–591.

Webb, L. W.: 1917. Transfer of Training and Retroaction. Psychological Review Monographs, 24: No. 104.

Wells, F. L.: 1908. Normal Performance in the Tapping Test. American Journal of Psychology, 19:437–483.

Wells, F. L.: 1912. The Relation of Practice to Individual Differences. American Journal of Psychology, 23:75–88.

West, A. F. (Editor): 1917. Value of the Classics, Princeton University Press.

Whipple, G. M.: 1910. Manual of Mental and Physical Tests.

Whipple, G. M.: 1916. How to Study Effectively. Public School Publishing Co., Bloomington, Ill.

Whitley, M. T.: 1911. An Empirical Study of Certain Tests for Individual Differences. Archives of Psychology, No. 19.

Wilbur, Flora: 1916. Experiments with Courtis' Practice Pads. Indiana University Studies.

Wilcox, M. J.: 1917. Does the Study of High School Latin improve High School English? School and Society, 6:58–60.

Wilson, G. M.: 1917. A Survey of the Social and Business Use of Arithmetic. Sixteenth Yearbook of the National Society for the Study of Education, 1917:128–142.

Wimmer, H.: 1916. An Experimental Study of the Effects of Drill in Arithmetical Processes under Varying Conditions. Indiana University Studies.

Winch, W. H.: 1904. Immediate Memory in School Children. British Journal of Psychology, 1:127–134.

Winch, W. H.: 1906. Immediate Memory in School Children: Auditory. British Journal of Psychology, 2:52–57.

Winch, W. H.: 1908. The Transfer of Improvement in Memory in School Children: I. British Journal of Psychology, 2:284–293.

Winch, W. H.: 1910. The Transfer of Improvement in Memory in School Children: II. British Journal of Psychology, 3:386–405.

Winch, W. H.: 1910. Accuracy in School Children. Does Improvement in Numerical Accuracy Transfer? Journal of Educational Psychology, 1:557–589.

Winch, W. H.: 1911. Further Work in Numerical Accuracy. Journal of Educational Psychology, 2:262–271.

Winship, A. E.: 1900. Jukes-Edwards: A Study in Education and Heredity.

Woodworth, R. S.: 1899. The Accuracy of Voluntary Movement. Psychological Review Monographs, No. 13.

Woodworth, R. S., and Thorndike, E. L.: 1901. The Influence of Improvement in One Mental Function upon Other Functions. Psychological Review, 8:247–261; 384–395; 553–564.

Woods, F. A.: 1906. Mental and Moral Heredity in Royalty.

Woody, C.: 1916. Measurements of Some Achievements in Arithmetic. Columbia University Contributions to Education, No. 80.

Yerkes, R. M., Bridges, J. M., and Hardwick, R. S.: 1915. A Point-Scale for Measuring Mental Ability.

INDEX