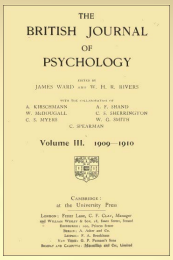
THE
BRITISH JOURNAL
OF
PSYCHOLOGY

Volume III. 1909—1910

British Journal of Psychology
Volume3, Issue1-2, December 1909

EXPERIMENTAL TESTS OF GENERAL INTELLIGENCE.

By CYRIL BURT,

Lecturer in Experimental Psychology, University of Liverpool.

(*From the Psychological Laboratory, Oxford.*)

I. *Origin and nature of the investigation.*

II. *Principles governing the selection of reagents and tests, and the application of the tests to the reagents.*

III. *Methods of provisionally estimating the intelligence of the subjects, and of mathematically deducing its relations to the tests.*

IV. *Apparatus, procedure, and results of the several tests* (*Touch, Weight, Sound, Lines; Tapping, Card-dealing; Card-sorting, Alphabet-sorting; Memory, Mirror, Spot Pattern and Dotting Apparatus*).

V. *Conclusions as to the possibility of the diagnosis, analysis and inheritance of intelligence.*

Origin and Nature of the Investigation.

The experimental determination of the mental characters of individuals is admittedly a problem of wide theoretical interest and of vast practical importance. The particular mental character which in importance is perhaps above all supreme, is that traditionally termed 'General Intelligence.' Yet the notice it has received from psychologists has been in proportion astonishingly scant. First suggested by Galton's *Inquiries* in 1883, first actually attempted upon some fifteen hundred American school-children in 1891, then assiduously prosecuted during the next ten years in America, in Germany and in France, eventually discredited by the hopelessly discordant results of the various researches, the investigation of Intelligence now seems once more to be attracting scientific attention. The success of anthropometry in statistically establishing relations between physical characters renders it not improbable that the discrepancies and failures of previous investigators of the relations between mental characters were largely due to their

reliance for the discovery of correlations upon mere inspection of the data they obtained, instead of upon quantitative determination and mathematical deduction. On this presumption, the mathematical methods of biometricians have recently been adapted and applied by Dr Spearman to psychological data obtained by himself, by co-workers, and by earlier pioneers, with positive and encouraging success. The experimental methods adopted for such investigations, the 'Test-Methods,' have been debated at length, and at last rehabilitated, in the latest and completest work on *Experimentelle Pädagogik*, by Professor Meumann, who had previously contributed some original experiments of his own. To the publications of these two writers reference may be made for a detailed history of the subject and for a summary of its literature[1].

The investigation reported in the following pages was commenced with a view to testing in practice the mathematical methods of Dr Spearman, and to verifying the experimental results both of Dr Spearman and of Prof. Meumann. For comparison with these some new experiments were also attempted, similar in aim but different in nature. The essential purpose of these was to determine whether higher mental functions would not show a yet closer connection with 'General Intelligence' than was shown by simpler mental functions, such as sensory discrimination and motor reaction, with which previous investigators had been so largely engrossed. The experiments were carried out—thanks to the extreme courtesy and kindness of Mr Claude Moore, Headmaster of the Central School, Oxford, and of Mr C. C. Lynam, Headmaster of the Oxford Preparatory School—mainly in the premises and upon the children of their respective schools during the autumn of 1907 and the spring of 1908. Throughout almost the whole period of the experimental part of the work I enjoyed the invaluable co-operation of Mr J. C. Flügel, of Balliol College, Oxford, whose results are here recorded with my own; while the mathematical part of the work is especially indebted to the generous advice and assistance of Dr Spearman. The whole research, however, owes its origin to the suggestion of Mr W. McDougall and its completion to his constant encouragement and advice. In conjunction with Mr Keatinge, Reader in Education in the University of Oxford, and Mr A. M. Hocart,

[1] C. Spearman, "General Intelligence Objectively Measured and Determined," *Amer. J. Psychol.* 1904, Vol. xv. pp. 202–292 (literature, p. 206 ff.). Krueger u. Spearman, "Die Korrelation zw. Verschiedenen geistigen Leistungsfähigkeiten," *Zeitschr. f. Psychol.* Bd. 44, S. 50 ff. E. Meumann, "Intelligenzprüfungen an Kindern der Volkschule," *Die Experimentelle Pädagogik*, 1905, 1 Bd. Heft 1/2 (literature *ad init.*). E. Meumann, *Experimentelle Pädagogik*, 1907–8, Vol. *i.* p. 386 ff. (Literaturverzeichnis, p. 552).

Mr McDougall had begun experiments on similar lines in his laboratory. Their results were not extensive enough for statistical treatment, but their experiments worked out some of the new tests to be described below, and proved their applicability to school-boys. The work of Mr Flügel and myself was thus considerably facilitated when we took up the investigation *de novo*. Both apparatus and rooms of the new Psychophysical Laboratory at Oxford were also lent most generously for our use. Dr Spearman and Mr McDougall have been kind enough to read through my paper in manuscript, and to allow me to make use of their criticisms and embody their suggestions.

To all those who in these and many other ways aided my work I take this opportunity of expressing my inadequate thanks.

Any attempt to elaborate a method for the determination of the presence of a given character in the mental constitution of individuals will at the same time throw light upon the nature and upon the development of that character. An investigation of 'General Intelligence' thus leads to three main enquiries:—(i) can its presence be detected and its amount measured? (ii) can its nature be isolated and its meaning analysed? (iii) is its development predominantly determined by environmental influence and individual acquisition, or is it rather dependent upon the inheritance of a racial character or family trait?

To decide between the possible answers to these three enquiries, or at least to contribute towards the definition of the problems involved and the methods available for their solution, seems a task of peculiar urgency for experimental psychology. For general psychology has not established the nature, nor individual psychology the measure, nor social psychology the transmission of 'General Intelligence.' The meaning of the term is assumed in pedagogical and sociological theory as generally understood; the property denoted by it is recognised in practical life by miscellaneous symptoms, such as physical and physiognomical characteristics, 'general impression,' or examination results. Yet the signs of intelligence are notoriously fallible, and its significance has either been ignored by introspective psychology, or else baffled its analysis. And so, with the current failure of theory in analysis, and of diagnosis in practice, the field is open for special experimental research. Whether Intelligence consists of a single elementary faculty; whether it is the complex resultant of a number of faculties, all working in co-operation; or whether there is really no such thing as 'General Intelligence'—the substantive being but the hypostatisation of an attribute applied to effects, apparently similar and practically equivalent,

but arising indifferently from a variety of alternative processes, which may operate independently on various occasions and in various individuals—these are controversies still awaiting the evidence of experiment. Further, if Intelligence consist only of some single, relatively simple function, such as adaptation of attention (suggested by Binet and others), or general sensory discrimination (suggested by Titchener and others), then Intelligence will presumably be recognisable by the success with which some simple task, demanding little but rapid adaption of attention or fine acuity of sensory discrimination, is performed; on the other hand, if it be a more composite function, then it will reveal itself in tasks involving mental process of a higher level and more complicated type; if a central or general faculty be a figment, if excellence in any one function or group of functions will serve as well as excellence in any other (as once suggested by Dürr[1]), then recognised intelligence will not associate itself uniformly with proficiency in one special direction, but will disclose itself, in some individuals by peculiar success in this or in that one particular direction, in other individuals by moderate success in several directions at once; but again the possibility and nature of any practicable tests can only be decided by means of experiment. Once devised, once demonstrated to measure a general, innate endowment, as distinguished from special knowledge and special dexterities, that is to say, from post-natal acquisition, such tests would find yet a third direction for experimental investigation, namely the enquiry how far the capacity thus measured varies with Age, with Education, with Parentage and with Social Rank; and this further application of the methods of the 'tests' would provide at once an illustration of their practical importance and a corroboration of their theoretical validity.

In attacking experimentally the several problems thus indicated the following course must be pursued: having selected suitable reagents and appropriate tests, to apply the tests selected to the selected reagents; to estimate the degree of correspondence between the results of the respective tests and the presumable intelligence of the reagents; to discover the highest common psychological factor that explains the various correspondences revealed; and to estimate the degree of correspondence between the native ability of the reagents—as

[1] *Zeitschr. f. Psychol.* 1906, Vol. XLI. April: "dass, was man im Leben 'Intelligenz' nennt, sich wissentschaftlich ebensogut als zufällige Verbindung einiger glücklicher Dispositionen verstehen lässt." Dürr, however, no longer endorses this view, and has retracted the chief criticisms founded thereon.

shown by those experiments that appear the better tests of unacquired intelligence—and the presumable intelligence of the reagents' parents. It is clear that a single set of experiments may thus be made to elucidate the threefold enquiry enunciated above.

The Selection of Tests.

The 'Tests' were psycho-physical tasks selected as readily and rapidly yielding comparable quantitative results, and as standing in possible functional relationship with the capacity called 'General Intelligence.'

To fulfil the first condition and thus subserve the practical diagnosis of Intelligence, it was necessary that the tasks should be fairly simple in character, dispensing as far as possible with elaborate or expensive apparatus; otherwise the time and cost entailed would preclude the subsequent application of the methods, if they proved promising, upon any extensive scale. For experiments upon young and untrained subjects such as ours, there is a further advantage in using none but the simplest apparatus. To boys, strange apparatus is distracting. Clock-work mechanism arouses irrelevant interests. Electric wires and keys inspire needless apprehensions. Consequently, in dispensing with elaborate instruments, the sacrifice of the mechanical regulation of objective conditions is often more than compensated by the exclusion of subjective irregularities and unstable attitudes of mind.

To fulfil the second condition, and thus contribute to the theoretical analysis of Intelligence, it was necessary that the tests should represent as far as possible the various main aspects and levels of mental process. The simplicity and practicability of the tests actually employed by us may be judged from the ensuing descriptions. Their representative character and limitations may be exhibited by the following classification (necessarily a somewhat arbitrary and schematic one) which attempts to arrange them according to the type of mental process predominantly involved in each.

List of Tests.	*Nature of process tested.*
I. Sensory Tests:	
(1) Discrimination of two points upon the skin	Perceptual discrimination.
(2) Discrimination of lifted weights	
(3) Discrimination of pitch	
(4) Comparison of length of lines by eye	

List of Tests.	*Nature of process tested.*
II. MOTOR TESTS:	
(5) Tapping (6) Card-dealing	Simple reactions.
III. SENSORI MOTOR TESTS:	
(7) Card-sorting	Reactions complicated by discrimination.
(8) Alphabet-finding	Reactions complicated by recognition.
IV. ASSOCIATION TESTS:	
(9) Immediate retention of (*a*) Concrete words (*b*) Abstract words (*c*) Nonsense syllables	Immediate memory.
(10) Mirror test	Formation of associations during motor activity (progressive process of 'Trial and Error').
(11) Spot pattern test	Formation of associations during perceptual activity (progressive process of 'Apperception').
V. TEST OF VOLUNTARY ATTENTION:	
(12) Dotting irregular dots	Maximal effort of sustained attention.

Of these, the Alphabet Test, the Mirror Test, the Spot Pattern Test, and the Dotting Test are believed to be, at any rate in this connection, new.

THE SELECTION OF REAGENTS.

The interdependence of mental processes is far more complex and far more intimate than that of physical processes. Thus the prime difficulty in a psychological research is the elimination of the factors that are irrelevant. In investigating General Intelligence by means of experimental tests the essential relations between the functions to be observed are liable to be distorted or obscured by such accidental conditions as the personality of the conductor of the experiments, the age, sex, social status, education, zeal and practice of the subjects of the experiments. Such sources of confusion may be reduced in three ways: their introduction may be evaded at the outset by circumspection in the choice of reagents; their interference may be nullified during the actual experiments by appropriate procedure and by repetition; their effects may be estimated in the results and discounted mathematically by subsequent calculation.

The differential influence of age, sex, education and social status may be minimised by selecting the groups of reagents from persons as far as possible similarly equipped in these several respects.

The reagents for the following investigations form two groups chosen from two Oxford schools,—a superior Elementary School and a high

class Preparatory School. Both were exclusively boys' schools. Difference of *Sex* was thus at once ruled out. In *Social Status*, the boys of the Elementary School were of the lower middle class, sons of local tradesmen, paying a fee of 9*d.* a week. The boys from the Preparatory School were being prepared for scholarships at one or other of the great Public Schools, and were in nearly every case sons of men of eminence in the intellectual world, that is to say, of Fellows of the Royal Society, University Professors, College Tutors and Bishops. Between the two schools there was thus a considerable difference of parentage. Such a difference should enable us to discover, not only if an application of the same tests to children of totally different types would reveal the same relations between the mental capacities tested and general intelligence, but also how far the tests were applicable to the problem of the inheritance of intelligence as well as to its analysis and diagnosis. Within the two schools, however, the social status of the boys was unusually uniform. Out of each school a further selection was eventually[1] made with a view to avoiding the serious influences arising from difference of *Age*. After several trials and special consideration it was decided to restrict the reagents to boys of a single year and to choose as limits the ages of 12 years 6 months and 13 years 6 months; older children had already come under the selective influence of superannuation regulations or scholarship examinations, while younger children were found to be scarcely equal to co-operating in prolonged and careful experimentation. In the Elementary School, 30 children fell within these age-limits; and in the Preparatory School, 13. For purposes of corroboration a dozen cases are sufficient (as will later be seen) to establish statistically a significant correlation; while within the limits feasible in the present investigation an increase of the main group of cases beyond about 30 would not advantageously diminish the probable errors. These 43 children, therefore, were selected as the final reagents.

Application of the Tests to the Reagents.

Irrelevant factors that could not be excluded in selecting the reagents,—such as *Personality*, both of operator and of subject,—we sought to exclude in selecting the methods to be adopted in the application of the tests. Before applying them to the selected reagents, most of the tests were first executed upon other boys taken either

[1] After the experiments recorded on p. 140.

from the same school or from a third school[1]. We were thus enabled to discover the errors incidental to the respective tests, to ascertain the more satisfactory modes of procedure with young subjects, and to habituate ourselves in the use of such procedure when adopted.

From the outset it was determined to examine the boys individually and personally rather than by class experiments or through the instrumentality of their teachers—thus differentiating our procedure from many, if not most, of our predecessors. Many of our tests could not have been carried out upon a number of subjects at once without needlessly multiplying apparatus and superintendents. Again, the writer had for a different purpose compared the results of tests carried out at one of these schools by mass-experiments with those of individual experiments; and there appeared no doubt that in the mass-experiments, even with the most rigid discipline, a number of undesirable factors intervened which could be ruled out or reduced in individual experiment. He ventures to think that high correlations obtained by the former method between imputed intelligence and particular tests—e.g. tests of visual and auditory discrimination—indicate that the more intelligent children will execute an unfamiliar, non-scholastic task (whatever its intrinsic nature) with the greater amount of interest, attention and care, rather than that their apparently higher thresholds measure their supposed greater acuteness in the mode of native sensory discrimination thus investigated, or that this acuteness is a component or a sample of general intelligence. Indeed, it might plausibly be argued that the complete theoretical correlation held to subsist between imputed intelligence and general sensory discrimination (i.e. the factor assumed to be common and fundamental to all sensory tests), really represents the correlation between imputed intelligence, and intelligence in the sense of power of attentively and carefully applying the mind to something relatively novel; for intelligence in this sense is inevitably manifested in all unfamiliar processes where much is left to the activity of the child; and may really have been the highest common factor in such tests.

The children whose performances are recorded in the sequel were therefore examined, not in class, but individually. They were also examined, not through the medium of their teachers or of other experimental conductors, but by ourselves personally. Here again the restricted numbers with which we had to cope enabled us to depart from the procedure of many of our predecessors. We found that, except

[1] We are much indebted to the Headmaster of St Philip's School, Oxford, for permitting us to perform some of these trial-series upon his pupils and at his school.

perhaps after prolonged collaboration, no two investigators adopt precisely the same procedure in tests where much depends upon the management of the operator; and that these variations in the details of procedure affect the measurement accepted as the result of the experiment, especially with children, to an undesirable extent. Since these and other consequences of the personal equation could not be completely avoided, it was resolved that as far as possible they should at least operate upon all alike.

In the actual course of the experiment the personality of the operator was relegated to the background by making our procedure as mechanical as could be, using a written form of instructions to be recited to each boy at the beginning of every new test. But since the same formula has not the same intelligibility or the same meaning for different children, this was always supplemented by asking if the reagent had any questions or difficulties as to the nature of his task. To make it possible to neutralise the personal factor by means of mathematical manipulation of the results nearly every test was applied, not only by the writer, but also by Mr Flügel.

Irregular influences arise not only from the personality of the operator, but also from that of the child. One of the most prolific sources of erroneous psychical measurement is difference of *Zeal.* The discipline of boys of a superior Elementary School and the good-nature of boys of a high class Preparatory School rendered our subjects surprisingly and uniformly good experimental reagents. Of the two groups, the Preparatory School boys were perhaps slightly superior in conscientious steadiness and care; while the Elementary School boys (who were markedly pleased at the interruptions of their regular routine, and were further fortified in their specially prolonged examination by the promise of a prize for the best) were perhaps slightly superior in spontaneous interest and attention. But in both groups attention, interest and goodwill were excellent, and consistently maintained.

Other sources of error lying in personal differences between the subjects were differences of *Practice* at similar tasks prior to serving as reagents, and differences of susceptibility to practice during the actual experimentation. The former affects only one or two of the tests applied, and will be noticed in connexion with these. The latter must vitiate in some degree every test no matter what may be its special nature. The ideal method of combating it would have been to insist on a maximum amount of practice in every case, to train each boy in each test at successive sittings, and accept only his final records as

furnishing the measurement required. This, however, was needless for the scope of the present enquiry, and would be impossible in any anthropometrical survey upon a large scale[1].

The customary procedure in other researches has been to adopt a medium and limited amount of practice, such as a "quarter of an hour's fore-exercise." In our preliminary trials we found that this had a highly differentiating effect upon the boys. Some reached their most favourable disposition almost at once, while the comprehension and interest of others were awakened much more gradually. In any case, a quarter of an hour at a task whose intrinsic interest for the reagent was small (such as a sensory-discrimination experiment) tends to prove tedious for him; and if immediately followed by the crucial investigation, which may last for another 20 minutes, the whole experiment usually resolves itself into a test either of his power to withstand boredom or of his tendency to get rapidly confused. Accordingly, with the exception of two tests involving elaborate apparatus, all express fore-exercise was, where possible, discarded; with the few boys, or on the few occasions, that needed special training, an interval for rest and recovery was allowed; otherwise only such training was given as was involved in the procedure of the actual test. Our procedure in regard to practice thus aimed at reducing it to a minimum.

The remaining factors likely to disturb the accuracy of the measurements, such as fatigue, variation in health, nature of the weather, time of day, and incidents of the occasion generally (all of which were found

[1] Those who see in tests of General Intelligence by psycho-physical experiment mainly a possible substitute for the present examination system commonly suggest that the chief and fatal objection to their introduction for practical purposes is that, once introduced, it would be impossible to prevent schoolmasters from training their pupils to their highest possible efficiency in the various tests just as they now tend to cram their pupils to their highest possible capacity with material for the examination, and that such special training would frustrate the intention of the new system just as it has frustrated the intention of the old. The defect of cramming in the examination system, however, is that it converts an instrument intended to test the subject's power of intelligently applying what he has learnt, into a mere test of the contents and retentiveness of his memory, and that these are very fallible symptoms of intelligence. The result of special practice in the experimental tests, on the other hand, would only be to convert a measurement of a special mental process on its first appearance into a measurement of that process at its maximum efficiency; and as this must form a far truer measurement of that process itself, it presumably would also form a superior test of the intelligence correlated with it. Simpler and quicker, such a course of special training would be far more generally available than a course of special cramming. And thus, by rendering the subjective conditions of the final test more constant and more universal, it would further and not frustrate their original design.

to exert a slight but observable influence), were more or less neutralised by repeated sittings. As has been mentioned, a second series was kindly undertaken for most of the tests by Mr Flügel. In the case of the more important tests a third series was again worked through by the writer, and in one or two instances a fourth and a fifth; but it appeared that little or nothing was gained in the average by more than three applications. At the Elementary School the repetitions were usually made at intervals of about a week, and at the Preparatory School at intervals of two days. Several tests were in execution during the same period, so that the same test might be applied to the various boys during the same portion of the same school "hour." But it was unfortunately impossible to ensure this in the case of all.

Both in the selection of reagents and in the application of the tests the principles of the present investigation have departed from those of previous ones. In endeavouring to abstract from accidental conditions, previous investigators of General Intelligence have commonly thrown together the performances of several hundreds of subjects, and trusted to numbers to neutralise in the average the manifold variety thus courted; or else they have taken haphazard a few available subjects, and relied on purely theoretical calculations for estimating and discounting the probable effects, whether of individuality or of occasion, that were irrelevant to the main issue. In the present investigation the subjects were as nearly as possible alike in all important respects save those investigated, and their number was small enough for each individual to be examined by the same operators and for each examination to be repeated at least once. To these features such positive results as the present investigation has attained are believed to be largely due.

Method of provisionally estimating Intelligence.

To determine the degree to which the various tasks might be considered satisfactory tests of General Intelligence, it was necessary to obtain an independent estimate of the relative intelligence of the reagents tested. For this, recourse was had at the conclusion of the experimental part of the work to their headmasters, their teachers, and their schoolfellows, who undertook to draw up on the basis of their general experience of the examinees independent lists, grading them in order of General Intelligence. No *a priori* assumption was made in the projection of this investigation as to what kind of mental capacity may with the greatest propriety be termed 'General Intelligence,' since it

is part of the aim of such investigations as the present empirically to examine the various capacities having a claim to this title and to ascertain their relations to one another. Hence, no determinate definition of 'General Intelligence' could be given to the compilers of the lists; rather it was presumed that the schoolmaster was the proper person, if any, to know the original meaning of intelligence, to recognise it in the concrete, and to compare its various degrees, even though the psychologist might prove the proper person subsequently to find for that meaning adequate expression, and to analyse and describe in technical terminology the nature of the capacity denoted by it. Accordingly, the serial classification of the children in order of intelligence, which serves as our provisional criterion in evaluating the experimental tests, was carried out by the compilers from a practical point of view with virtually no interference whatever from the experimenters in their capacity as psychological theorists[1]. And, of course, during the progress of the compilation and of the experiments, both compilers and experimenters remained in ignorance of the lists of each other.

The actual procedure of the compilers in arranging the names in the required order seems to have been somewhat as follows:

In the case of the children from the Elementary School, the Headmaster made three lists of the 30 boys who had served as subjects in these experiments, according to the class-lists of the three standards to which they respectively belonged; these three lists he connected into one by carefully intercalating the bottom boys of the upper standards with the top boys of the lower standards. He then thoroughly scrutinised the order and further re-arranged it from his private knowledge of the boys, with each of whom he was personally familiar. After an interval of several weeks, during which he frequently took lessons with the standards in question, he again revised the list. Where in doubt as to the relative position of two or more boys, his test-question was: "Which boy is the quickest at seeing the point of anything?" From his reputation as a judge of character, from his long personal experience of the boys concerned, and from the special interest, care, and conscientiousness with which he performed the task, there can be little doubt that the grading is as nearly perfect as a grading based on personal impression could be.

[1] It may be remarked that all who were kind enough (at the cost of considerable trouble, time, and care) to draw up these lists recognised, in one way or another and in their own phraseology, the distinction between successful mental activity due to sheer intelligence, independent of special experience of the particular subject-matter of that activity, and successful mental activity due to "learning" in one form or another, i.e. to acquired knowledge or to habitual practice; and were reminded that the preservation of this distinction was the most important and the most difficult part of their task.

For the mathematical elimination of observational errors by means of the formula presently to be cited, it was necessary to have at least two gradings for the same capacity, as far as possible independently obtained. A second grading of the boys, therefore, was obtained from their class masters. A third grading was furnished by their schoolfellows,—two competent and impartial boys, not themselves among the thirty, being selected for the purpose. The boys were asked: 'supposing you had to choose a leader for an expedition into an unknown country, which of these 30 boys would you select as the most intelligent? Failing him, which next?' And so on.

In the case of the 13 boys from the Preparatory School, the Headmaster's estimation was again based on the class-orders. By revising these, he produced two lists in order of Literary and Mathematical Ability respectively; and from an amalgamation of these a final order of general intelligence was derived. A supplementary grading for the purpose of mathematical correction was also obtained from two boys holding responsible positions in the school, not themselves on the list.

The Method of Calculating the Correlations.

The tendency to concomitant variation between two mental characteristics, such as General Intelligence and proficiency in some experimental test, may be best expressed by means of a coefficient of correlation. A coefficient of correlation is a single figure so calculated from a number of individual measurements as to represent with quantitative precision that degree of relationship between two variable qualities of a group, from which all the measurements actually observed might have arisen with least improbability. A relationship may be either absent, or present; and if present, either negative (i.e. inverse), or positive; and either complete, or more or less incomplete. Accordingly, the coefficient expressing it may have all possible values, from -1 through 0 to 1; or, in percentage, from -100 °/₀ through 0 °/₀ to 100 °/₀.

A method of calculating correlational coefficients, convenient and of sufficient accuracy for the present purpose, has been devised, and used in investigations similar to the present, by Dr Spearman. It is called by him the 'foot-rule' or R-method. This method is a simplified application of the standard, 'product moments,' or r-method, elaborated by Bravais, Galton and Pearson. The actual measurements of the reagents' thresholds (or other capacities) forming the two series

to be compared are first converted into comparable, i.e. homogeneous, values by substituting for them the numbers denoting the relative rank or position of the several reagents as arranged in order of proficiency for the two tasks.

The sum of the discrepancies in rank between the two series is then found. This is most readily done by subtracting the number representing the position in rank of each individual for the second test from the number representing his position in the rank for the first test, wherever the former is smaller; the remainders, representing the number of places gained by the individuals better at the second test, are added together, and multiplied by two; for the number of places gained by these individuals must necessarily be the same as the number of places lost by the rest, who are worse at the second test; and thus the sum of discrepancies between the two tests will be equal to twice the sum of gains in the second test.

The sum of the discrepancies between two series of the same length to be expected on an average by mere chance is then to be obtained by squaring the number of the members, subtracting 1 from the square, and then dividing the remainder by 3.

The ratio is next found between the sum of the discrepancies in rank between the two series observed and the sum of rank discrepancies to be expected by mere chance; and the resultant fraction is subtracted from unity. This gives the coefficient for the correlation between the two series in terms of the 'foot-rule'[1]. (For illustration and proof, see Spearman, this *Journal*, *loc. cit.* p. 107.) In algebraic notation, if R denote the coefficient of correlation, Σd the observed sum of dis-

[1] C. Spearman, *Brit. Journ. of Psych.* 1906, Vol. II. Pt. 1, p. 89, "'Foot-rule' for Measuring Correlation." Cf. also *ibid. Am. Journ. of Psych.* xv. 1904, pp. 72-101, "The Proof and Measurement of Association between Two Things"; also *id. ibid.* pp. 202-292, "General Intelligence Objectively Determined and Measured." The validity of the formulae given in the *American Journal* was challenged by Karl Pearson in *Biometrika*, III. p. 160; detailed mathematical proofs, however, have since been furnished, *Am. J. Ps.* XVIII. p. 161, and *Brit. J. Ps. l.c.* A critical discussion of the Foot-rule method by Karl Pearson will be found in *Biometric Series*, vol. II. (Drapers' Company Research Publications), and of the Correction formulae by William Brown in the reports of the *VIme Congrès Internationale de Psychologie, Genève*, 1909 ('Some Experimental Results in Correlation'). In this article I have nowhere discussed at length the advantages and limitations of Dr Spearman's formulae: instead of burdening a lengthy paper with further mathematical technicalities, I have preferred to leave the merits and defects of the methods I have endeavoured to test and illustrate to be inferred from the results they yield by those more competent to form an opinion on mathematical questions than myself. I may, however, add that most of the figures essential to my general psychological conclusions have been checked by the more elaborate and better accredited formulae, and the discrepancies have been found to be practically negligible.

crepancies, Σg the observed sum of gains, and n the number of cases in each series, then

$$R = 1 - \frac{\Sigma d}{\frac{n^2 - 1}{3}} = 1 - \frac{\Sigma g}{\frac{n^2 - 1}{6}}.$$

Thus, when there is complete correspondence between the two series of measurements and consequently no discrepancies of rank, the correlational coefficient will obviously be 1; when there is no correspondence the number of discrepancies will presumably be approximately the same as the number of discrepancies expected by pure chance, and the coefficient will consequently be 0 or nearly 0. When there is complete inverse proportionality the coefficient should be -1. This is actually the case when the coefficients are calculated by the standard or r-method, the method of product moments. But when calculated by the foot-rule or R-method the positive values obtained, especially midway between 0 and $+1$, are smaller than those obtained by the standard method, and the negative values are all larger. An inverse or negative correlation can usually be evaded by reversing the ranks of one of the series compared; while the positive values for R can readily be converted into terms of r by the empirical table given by Spearman (*B. J. P.*, *l. c.*, p. 104)[1]. On account of its convenience the R-method has been preferred in the present work for calculations, and it has not been thought necessary to evade by re-calculation the exaggeration of negative values. But as coefficients obtained by the standard method are more familiar to statisticians and admit more readily of the application of correction formulae, the translation from 'foot-rule' values has always been made, and the figures printed represent throughout correlations in terms of r.

In determining the relationship between two qualities from their manifestations in a limited group or sample, instead of from a complete investigation of the entire class—an investigation we never could accomplish—we are admitting a determinable source of error. The expression representing this source of error is called the Probable Error. It must obviously decrease in some way with the increase of the number of cases employed. For the correlations given below the probable error may, legitimately enough for the present purpose, be calculated directly from the more complete formula of the r-method; since the probable errors of the two respective methods are approximately the same in size,

[1] Or by the formula since mathematically deduced by Prof. Pearson (*Biometric Series*, *l.c.*). The average discrepancy, however, between the values given by Prof. Pearson's formula, and those given by Dr Spearman amount to less than 0·01.

(On the existing formulae for the probable error for coefficients obtained by Spearman's methods, see Karl Pearson, *loc. cit.*)

and the coefficients are everywhere converted into terms of r[1]. The formula for the r or standard method is taken as

$$\text{p.e.} = \frac{\cdot 6745}{\sqrt{n}}(1 - r^2).$$

Thus, if $r = \cdot 50$,—where $n = 30$ (number in Elementary School group), $\text{p.e.} = \frac{\cdot 6745}{\sqrt{30}}(1 - \cdot 50^2) = \cdot 09$; where $n =$ only 13 (number in Preparatory School group), $\text{p.e.} = \frac{\cdot 6745}{\sqrt{13}}(1 - \cdot 50^2) = \cdot 14$. A coefficient of correlation has little or no significance unless it is at least two to five times as great as its probable error. A coefficient five times as great as the probable error occurs by chance only once in 1000 trials; accordingly where a high correlation, such for instance as would give a coefficient $r = .50$, obtains between two functions, its existence may be satisfactorily demonstrated by about a dozen cases; and below this number none of the groups here investigated falls. A coefficient only twice as large as the probable error occurs about once in six times by mere chance. Hence such small coefficients can but suggest, not prove, the existence of real correspondences.

Besides those errors inevitable to the process of sampling and expressed by the 'probable error,' there are others incidental to the

[1] Dr Spearman has furnished me with the following proof that the probable errors of the two correlational methods are about equal:

Let R_R denote the p.e., expressed in terms of R, when using the R-method. Then

$$R_R = \frac{\cdot 43}{\sqrt{n}}. \qquad (\textit{Brit. J. Psychol.}\ \text{Vol. II. p. 106.})$$

Next let r_r denote the p.e., expressed in terms of r, when using the r-method. Then

$$r_r = \frac{\cdot 6745}{\sqrt{n}}(1 - r^2).$$

But it has been shown (*Brit. J. Psychol.* Vol. II. p. 102) that

$$r_r = \sin\left(\frac{\pi}{2} \cdot R_r\right)$$

(where R_r is the value corresponding to r_r in terms of the R-method); the latter expression, $\sin\left(\frac{\pi}{2} \cdot R_r\right)$, when, as here, small, is known to be approximately equal to $\frac{\pi}{2} \cdot R_r$.

Hence $$r_r = \frac{\pi}{2} \cdot R_r,$$

so that $$R_r = \frac{\cdot 6745}{\sqrt{n}}(1 - r^2)\frac{2}{\pi};$$

and, neglecting r^2, $$= \frac{\cdot 6745}{\sqrt{n}} \cdot \frac{2}{\pi} = \frac{\cdot 429}{\sqrt{n}} = \text{very approximately } R_R.$$

determination of correlations which are also susceptible of mathematical treatment. Errors traceable to definite and controllable sources have been largely minimised empirically by the principles governing the selection of reagents and the application of the tests; the remainders of these, together with errors creeping in from miscellaneous uncontrollable sources, constitute Observational Errors; such are inevitable to all measurement. All observations are liable to fluctuations, due to the fact that our figures measure not purely and directly the quantity they are intended to measure, but that quantity as modified by irregularly varying observational conditions; and psychological observations with children are peculiarly difficult and indirect. But upon repetition of the measurements, and appropriate calculation, such fluctuations may either neutralise each other in the average, or themselves be estimated and corrected in the final result.

As has been stated, two or three series of applications were made with nearly every test. The simplest method of eliminating observational error is to find the average of the two or three measurements actually observed for each individual, and work out the correlation between the two series of averages. The coefficient thus obtained is called the Coefficient of the Amalgamated Series.

But in so doing we have clearly sacrificed an additional datum supplied by the actual figures as experimentally obtained, namely, the amount of variability of these original figures about the calculated average. A set of averages derived from figures showing but small fluctuations, and consequently involving but slight errors of measurement, is more reliable than a set of averages derived from figures exhibiting large fluctuations and thus involving relatively large errors of measurement; such wide deviations and such unreliable averages must evidently tend to obscure a correlation when present and to diminish its apparent coefficient. The amount of fluctuation to which a given mode of measurement is liable can be determined by finding the correlation between the two sets of experimental figures obtained on different occasions or by different observers for one and the same measurement, or the average correlation if the measurement was applied more than twice. The coefficient then obtained is called the Reliability Coefficient, and measures inversely the reduction of apparent correlation caused by errors of observation incidental to that measurement. To eliminate such illicit reduction in the correlation between two such modes of measurement, the arithmetical mean of the several correlations between each of the various series repeated with one mode

of measurement, with each of the various series repeated with the other, is divided by the geometrical mean of the two reliability coefficients for each of the two modes of measurement. Take for instance the determination of the true amount of correspondence between some mental test on the one hand and the estimation of intelligence by teachers on the other. Let us suppose that the test has been applied several times to a group of school children, and that these children have also been ranked according to intelligence by several school teachers. Let us denote the average reliability correlation between the several series of applications of this one particular test by $r(TT)$; the average reliability correlation between the orders for intelligence drawn up by the different teachers by $r(II)$; the average crude correlation of each of the series for the test with each of the orders for intelligence by $r(TI)$; and, finally, the required Corrected Coefficient of pure correlation between capacity tested and Intelligence by $r'(TI)$: then[1]

$$r'(TI) = \frac{r(TI)}{\sqrt{r(TT).r(II)}}.$$

When the coefficient of correlation is raised by such correction the probable error of the coefficient will necessarily rise as well. A fair approximation for the probable error of the corrected coefficient may be gained by the aid of the probable error of the amalgamated series and the use of the following formula:

$$\frac{\text{p. e. of corrected coefficient}}{\text{p. e. of coefficient of amalgamated series}}.$$

$$= \frac{\text{corrected coefficient}}{\text{coefficient of amalgamated series}}.$$

Suppose, for instance, that the coefficient of the amalgamated series comes to ·48, and the corrected coefficient to ·55, and that the number of members in the series correlated is 30, then the probable error of the amalgamated series

$$= \frac{\cdot 6745}{\sqrt{30}}(1 - \cdot 48^2) = \cdot 096;$$

hence the probable error of the corrected coefficient

$$= \cdot 096 \times \frac{\cdot 55}{\cdot 48} = \cdot 11.$$

[1] C. Spearman, *Am. J. Psychol.* 1907, Vol. XVIII. p. 161. See, however, criticisms by W. Brown, *loc. cit. sup.*, and comments below pp. 137, 160.

From this it is evident that nothing but an illusory result can be obtained by applying a big correction to a correlation having a big probable error. Mathematical correction can only be employed to deduce theoretically the probable amount of the pure correlation when the size of the crude coefficient, being at least twice its probable error, indicates that the observed correlation actually exists.

Non-Experimental Correlations for Intelligence.

The correlational results for the empirical estimation, furnished by masters and boys, of the reagents' relative intelligence may be given first. The reliability coefficient for the Elementary School group is ·88; for the Preparatory School group ·91. The estimates made by the Headmasters, especially for the larger group, were without doubt unusually reliable. Nevertheless the reliability coefficients are probably a little too high. This may, perhaps, be due to the fact that all who drew up the lists in question were more or less aware of the official class order, and may consequently have been influenced consciously or unconsciously by its arrangement, especially perhaps where their own personal impressions were deficient. If so, their gradings would to that extent fail to represent two series of observations completely independent.

No exact scholastic data were obtained to supply an absolute, as well as a relative, empirical evaluation of our boys' level of intelligence. At the Elementary School I was kindly allowed to set a special examination to the boys in Arithmetic, Composition, and other school subjects; but it did not prove practicable to set a similar examination to the boys of the Preparatory School, the results of which should be comparable to those of the former. The examinations referred to in the sequel, therefore, are the annual school examinations. The correlations of these with intelligence, as estimated by the masters' impressions, are ·81 at the Elementary School, and ·78 at the Preparatory School. The reliability coefficients of such examinations is between ·60 and ·80.

Though irrelevant to the present investigation, the following miscellaneous correlations obtained from the Elementary School group during our work there may be of interest to some. With them are printed, for purposes of comparison, coefficients obtained by Prof. Karl Pearson from a large number of schools, mainly secondary, taken from *Biometrika*, 1906–7, p. 127:

	Oxford Central School	Secondary and other Schools
Intelligence (Headmasters' estimate) and fairness of hair ...	·32	·10
Health (Masters' estimate)	·24	·17
Athletics ,, ...	·37	·20
Conscientiousness (Masters' estimate)	·54	·46
Age	·28	·05
Handwriting	·41	·28
Suggestibility	·28	
Repetition (Special Examination, 20 to 26 boys only)	·51	
Arithmetic ,, ,, ,, ...	·62	
Mental Arithmetic ,, ,, ,, ...	·73	
Dictation ,, ,, ,, ...	·77	
Composition ,, ,, ,, ...	·79	
Amalgamated result ,, ,, ,, ...	·74	

Suggestibility (Binet's line test) correlates with Conscientiousness to the extent of ·30. With reference to the supposed racial character of fair hair it may be added that over 75 °/₀ of the Preparatory group, and under 25 °/₀ of the Elementary group, appeared typically fair.

NATURE AND RESULTS OF THE SEVERAL TESTS.

At the risk of tediousness the apparatus and methods employed for the various tests are described with some detail, in order to enable anyone from the mere description to repeat the experiments with sufficient similarity for comparison with the results here recorded, or with the improvements suggested by our own experience for the achievements of results more accurate still.

The results—reliability coefficients, coefficients of amalgamated series, average raw coefficients, and corrected coefficients, together with the observational data in summary form—are first recorded and discussed under the headings of several tests; a comparative survey of the entire series may be obtained from the tables given later with the general conclusions.

For the sake of brevity, the observational data are recorded not in the form of the actual measurements experimentally obtained, but in the form of averages for each of the groups (calculated from the average performances for each of the members)—the distribution of the measurements within the groups being indicated by the mean variation of the

performances of the individuals[1], and their range by the two extreme measurements, i.e. by the average performances of the best and worst individuals in each particular test. A comparison of the respective data for the two groups will yield a conception of the capacities—both special and general—of the boys of superior parentage relative to that of the boys of ordinary parentage. The boys of the larger group, when arranged in order of intelligence by the Headmaster, seemed to him to fall naturally into three sections, comprising 7 "clever" boys, 14 "average" boys, and 9 boys "below the normal"; the average performances for each of these three sections is, accordingly, also given, since they not only afford a simple corroboration of correlations with intelligence, but also render the comparison between the two groups more definite. At the same school there was found to be a boy congenitally weak-minded; his age at the time of the commencement of the investigation was 13 years 10 months; he, therefore, did not fall within the group selected, but where it proved possible to apply the tests to him the results are recorded, since his intelligence was presumably minimal, and his case therefore furnishes a useful negative instance.

Sensory Tests.

(1) Touch Discrimination.

Apparatus and Procedure.

Tactile acuity was measured by the threshold for simultaneous discrimination of two points upon the skin. The area of skin tested was the middle third of the volar surface of the right forearm. The instrument used was the aesthesiometer devised and used by Dr Spearman[2]. Being constructed of aluminium with blunt celluloid tips, it is peculiarly light (25 grs.). It carries a millimeter scale and vernier for adjusting the distance between the two parallel points, and a third point for single stimulation. In preliminary trials upon other subjects the compasses were also employed, and found practicable, but not so convenient; the thresholds given were similar, but more variable.

[1] Designated, in accordance with Rivers' suggestion (*Brit. J. Psychol.* Vol. I. Pt. 4, pp. 354, 355) by the capital letters M. V., to distinguish it from the mean variation as a measure of the variability of observations designated by small letters, m. v.

[2] C. Spearman, *German Congress for Experimental Psychology*, 1904. K. Krueger u. C. Spearman, "Die Korrelation zw. verschiedenen geistigen Leistungsfähigkeiten," *Zeitschr. f. Psych.* Bd. 44, S. 69.

Each boy was allowed to watch and to emulate the performance of his predecessor. If it was his first occasion, the nature of his task was also briefly demonstrated to him. And then, the boy being seated, with his head averted, and his arm lying comfortably on a small table, screened from surreptitious glances, the points of the instrument were applied to the skin with firm pressure lasting nearly a second. It was found advisable to warn the subject by saying "Now" about one second before each application. The single point was applied almost as frequently as the two points. The subject was told to reply "One" or "Two" according as he judged that one or two points touched his arm. And his judgments were recorded (where possible) by a third person.

The threshold sought was the smallest distance at which two points yield a sensation perceptibly double, in order as far as possible to avoid the influence of inference from secondary sensory factors other than distinct twoness[1]. To take as the threshold the smallest distance at which two points yield a sensation merely different from that yielded by a single point, tends (unless all the reagents are first subjected to prolonged training) to handicap favourably those boys whom Binet[2] would class as *interpretateurs* at the expense of the *simplistes* and the so-called *distraits*; such a classification of subjects according to their different methods of judgment in this test was found in preliminary experiments to correlate with intelligence. The procedure followed was a combination of the Method of 'Minimal Changes' with the Method of 'Right and Wrong Cases'—a procedure devised and adopted by Mr McDougall for the investigation of the delicacy of Tactile Discrimination among the Murray Islanders[3].

Each series was commenced with applications of the two points at a distance well above the probable threshold of the subject (usually about 70 mm.), and proceeded by gradually diminishing this distance by successive stages till it had descended well below the threshold. So long as the boy's answers evinced no genuine errors, five applications of the two points were made at each stage, the interval between them was

[1] On this cf. especially W. H. R. Rivers, *Brit. Journ. of Psych.* Vol. I. Pt. 4, pp. 366 sq., 364 and 390.

[2] *L'Année Psychologique*, 1897, III. p. 225. In the two groups whose achievements are recorded below, there were noticed but two *interpretateurs*, and but three or four *distraits* (owing to imagination or suggestibility rather than inattention),—all at the Elementary school.

[3] W. McDougall, *Cambridge Anthropological Expedition to Torres Straits*, Vol. II. Pt. 2, p. 189.

diminished by 5 mm., and the boy was told whether his answers were right or wrong; this modification of the usual procedure was eventually adopted as the quickest method of leading the boy to feel confidence in his own judgments without developing and practising a tendency to speculate and infer. As soon as the boy's answers exhibited a genuine error, ten applications of the two points were given at each stage, and the distance was reduced by 2·5 mm., till a large percentage of errors was obtained. The stage at which the boy first began to make 20°/₀ errors, i.e. two wrong judgments in ten applications of the double[1] stimulus, was accepted (unless subsequent recovery in the continuation of the experiment indicated temporary inattention) as his threshold of tactile discrimination, and the distance in millimeters of the two points at this stage was accepted as the measure of his threshold. It may perhaps be argued that strictly thresholds are only comparable when, not the actual measurement of absolute distance just discriminable, but the ratio of this distance to the length of the arm of the subject, is taken as the measure of the threshold[2].

Results.

Three series of observations upon Tactile Discrimination were obtained from the 30 boys of the Elementary School, by the writer, by Mr Flügel, and by the writer again, at intervals of about a fortnight. Each observation occupied about 20 minutes, and consequently each series extended over several days. The influence during this period of changes of weather and of health, of the time of the day and of the week, was noticeable, but not considerable; and when any extraneous factor appeared to be influencing the work of a particular boy to a marked degree, manifesting itself in a suspiciously high threshold or (more commonly) in suspiciously irregular judgments, the sitting was adjourned. Two series of observations were obtained from the 13 boys at the Preparatory School, one by each of us, at an interval of four days. The number of subjects being here far smaller, variations of weather, fatigue, &c. were unmarked during the relatively short duration of the

[1] According to Mr McDougall's original procedure. Dr Rivers in adopting it prefers to use the errors in the single stimulations, not merely as a rough precaution and guide, but also as indicating the proper estimate of the threshold, which he defines as that distance at which two mistakes in ten occur with *each* kind of stimulation (*loc. cit.* p. 364).

[2] The forearms of the boys of the Elementary School varied considerably in length; consequently their thresholds would perhaps be more strictly comparable if they were first reduced to terms of an arm of average length. When this is done the correlation is slightly diminished, viz. to ·08.

observations. The reliability coefficient for the three Elementary School series was ·73; for the two Preparatory School series, ·75. The lower coefficient in the larger group may be due to the peculiar variability commonly evinced in sensory tests by about 20 °/₀ of any group of subjects, and actually traceable in the experimental figures yielded by this group, since for the above reasons this variability probably had a slightly freer opportunity to manifest itself at that school. But a difference of ·02 is in itself not a significant one. Excepting these characteristically variable subjects, and considering what at the time of experimenting we took to be the peculiar difficulties of applying sensory tests to school-children—the thresholds independently arrived at prove in the majority of cases to be strikingly uniform. Nor is there any significant difference between the reliability coefficients of two series obtained by two different operators and those of two series obtained by the same operator.

In an investigation into the *Korrelation zwischen verschiedenen geistigen Leistungsfähigkeiten* by Krueger and Spearman the reliability coefficient between the two series obtained by the two different operators upon eleven German subjects tested with similar apparatus upon the right hand, and the right and left cheeks, is given as only ·42. From this we may perhaps infer the relative superiority of the psycho-physical method devised by Mr McDougall and employed in the present investigations to the more rapid psycho-physical procedure employed by Krueger and Spearman.

The coefficients of correlation between the amalgamated series of average individual thresholds and the orders of Intelligence given by the Headmasters are for the two schools ·13[1] (p.e. ·12) and −·06 (p.e. ·20) respectively. The average of the uncorrected correlations between the several original series and the various gradings for Intelligence is ·14 for the Elementary children and −·14 for the Preparatory children. Corrected by the above reliability coefficients these would give pure correlations of ·18 and −·17; correction, however, would here be scarcely valid, as the average raw correlations are not twice the probable errors.

The correlations obtained by Krueger and Spearman between tactile thresholds and tests which we may regard as indicating the relative intelligence of their subjects—adding figures, Ebbinghaus's *Kombinations-Methode*, and learning rows of numbers—were correspondingly small, namely, 0·19, 0·00, and −0·13 (Probable Error,

[1] *Loc. cit.* S. 77.

0·19 to 0·25)[1]. Many observers have argued for a close connection between tactile discrimination and intelligence. Of these the most recent is M. Schuyten. In his experiments upon the cheeks of Antwerp school children, he finds that the intelligent half of each group tested has a decidedly acuter average threshold than the unintelligent half[2]. With him, however, "les expériences, une minute par élève," were extremely short. Such rapid methods with the aesthesiometer may yield (and apparently often have yielded) a correlation with intelligence; but it cannot yield a valid measure of sensory discrimination.

We may accordingly conclude that a genuine threshold for successive discrimination of two points upon the skin has extremely little or no connection with General Intelligence. This is corroborated by the observation that the congenital imbecile seemed to have the best threshold of any examined, namely less than 30 mm. according to one observer (Burt) and 10 mm. according to another observer (Flügel), even though our psycho-physical procedure was particularly unfavourable to such a subject. Though too small as compared with its probable error to be significant by itself, the positive correlation in the larger group accords with the fact that the three sections comprising clever, average and infra-normal boys had average thresholds of 31·3, 37·3, and 38·3 mm. respectively[3]. But this small correlation, if really existent at all, is probably to be attributed to the more intelligent attitude adopted by the more intelligent boy in submitting himself to the test, and to his consequent smaller percentage of accidentally erroneous judgments and higher percentage of correct inferences—a difference among the subjects which was scarcely observable among the Preparatory boys.

The observational data for comparing the Tactile Acuity of the two groups are as follows. Average threshold for discrimination of two points upon the forearm of boys tested at the Elementary School 36·2 mm. (Mean Variation 9·0, lowest threshold 19 mm., highest threshold 58·3 mm.); at the Preparatory School 38·9 mm. (Mean Variation 11·0, lowest 12·5 mm., highest 63·7 mm.). The average threshold at the latter being actually duller than the average threshold of the dullest group at the former (38·3 mm.), it is clear that sensory discrimination upon the skin is acuter at the Elementary School. Perhaps this may be plausibly attributed to an endowment making for readier and sharper sense-perceptions and perceptual 'inferences,' currently credited to the children of lower social status as compared on the

[1] *Loc. cit.* S. 75.

[2] *Revue de Psychiatrie*, 1908, p. 135.

[3] Or, reduced to terms of a forearm of average length, 31, 37, and 37 mm.

whole with the steady and more cautious child of the thoughtful classes. Confirmation of this suggestion, and of the data upon which it is based, may be found in measurements recently recorded by other observers. Yasousabouro Sakaki has tested the tactile discrimination upon the cheek of children at Tokio. He finds "la normale physiologique est pour les écoles primaires de filles 11,6 mm., pour les écoles primaires de garçons 12,3; pour les écoles supérieures de jeunes filles 12,1; pour un athenée 13,2." And again, he finds as the mean threshold of sons of 'petit négociants' 12,2 mm.; of sons of 'savants, professeurs' 13,0 (that of sons of 'banquiers,' however, is yet higher, namely 13,4)[1]. The figures for schools of different social status may be paralleled by figures for races of different culture. Arranging the available figures[2] for thresholds upon the forearm in order of magnitude, commencing with thresholds of greatest acuteness, we have the following series:

TABLE I.

Tactile Discrimination among Groups of Different Cultural Status.

Group	Observer	Threshold
Papuan boys	McDougall	15·0 mm.
Papuan men	McDougall	19·8 mm.
Dayaks	McDougall	35·0 mm.
Toda boys	Rivers	35·0 mm.
English Elementary boys	Flügel and Burt	36·2 mm.
English Preparatory boys	Flügel and Burt	38·9 mm.
English village children	Rivers	43·0 mm.
English men ('mostly of working classes')	McDougall	44·6 mm.
English men ('of the educated class') ...	McDougall	'rather higher'
Toda men	Rivers	45·5 mm.
English men (Cambridge graduates and undergraduates)	Rivers	56·5 mm.

In obtaining these figures, the procedures used by Mr McDougall, Mr Flügel and myself were the same; that used by Dr Rivers yields (as has been noted above) thresholds slightly higher than the procedure suggested by Mr McDougall. If, therefore, we made allowance for this difference, the series would rearrange itself, and the groups would fall in an order showing a complete inverse correspondence with that of

[1] *Revue de Psychiatrie*, 1908, pp. 140, 142.

[2] W. McDougall, *Anthropological Expedition to Torres Straits*, Vol. II. pp. 191, 192; W. H. R. Rivers, "Observations on the Senses of the Todas," *Brit. Journ. Psychol.* 1905, Vol. I. Part 4, p. 369.

cultural development. For, as Dr Rivers observes "in order of cultural development the Dayaks and Todas occupy an intermediate position" between Papuans and English; the Todas, I understand, are perhaps a little higher than the Dayaks; while the respective culture of the various English groups is implied in their description.

This, then, confirms our observations, that among groups differing considerably in cultural status, and so far presumably of different innate intelligence, the less intellectual group tend in average tactile discrimination to be the more acute; while among individuals of the same cultural class, any apparent positive correlation between tactile discrimination and intelligence, is probably illusory, and due to the interpretative quickness of the more intelligent.

(2) Weight Discrimination.

Apparatus and Procedure.

Capacity to discriminate Weights was investigated by means of a graduated series of some 21 weights, constructed by the Cambridge Instrument Company on Galton's convenient cartridge pattern, ranging in weight from 80 to 120 grams. These were lifted by the subject with the thumb and finger of the right hand through a vertical distance of 17·5 cm.—the height of the lift being regulated by a tape stretched horizontally between two uprights. The standard, a constant[1] weight of 100 gms., was lifted first; the variable weight, placed on the right of the standard weight, was lifted second. And the boy was instructed to say, after once simply raising and replacing each of the pair of weights set before him, whether the second weight was 'lighter' or 'heavier[2].'

The psycho-physical method adopted was analogous to that described as used in testing discrimination of two points. As before a 'descending series' alone was taken, and at the crucial stages the procedure was

[1] To ensure that the judgments were immediate, and not illegitimate inferences based on surmises as to the nature of the psycho-physical method (e.g. when there was reason to suspect that a boy recognised the variable weight as the same during its repetition at a given stage from accidental marks upon the cartridge cases), it was sometimes necessary to alter the standard slightly, and make a corresponding alteration in the variable weights.

[2] It may be of interest to note, as bearing on the psychological theory of Comparison of Sense-impressions, that the natural tendency of the boys seemed invariably to be to indicate, by pointing or by naming, the heavier of the two weights rather than to pronounce a judgment directly expressing an 'absolute impression' of the heaviness or lightness of the last lifted.

'without knowledge' of the correctness of the judgment. The subject commenced with an easily distinguishable positive difference, viz. 18 gms. (100 gms. standard and 118 gms. variable), and the difference was reduced by successive steps of 2 gms. An equal number of equal negative differences, i.e. differences below the standard, were irregularly interspersed at each step to obviate a bias in favour of the judgment that the second was generally heavier. The point at which the subject first made 20 °/₀ errors upon the positive differences without subsequent recovery was accepted as his threshold.

Results.

In successive discrimination of lifted weights, three experimental series were carried out at the Elementary School and two at the Preparatory School, as for Touch. At the Elementary School the determination of the threshold for weight gave far less trouble to both subjects and operators than that for touch. But by nearly all the Preparatory boys it was much more confusedly performed. The reason is difficult to assign; observation of their manner when judging suggested that they found greater difficulty in associating difference of mass (unless very distinct) with objects of equal size and similar appearance. The fact reappears conclusively in the reliability coefficients; these are for the one school ·86, and for the other ·51.

The correlation between the average thresholds and the Headmasters' grading is – ·13 (p. e. ·12) and – ·19 (p. e. ·19); the average crude correlation – ·01 and – ·14; the corrected coefficients would work out at – ·01 and – ·20, but correction is again scarcely valid.

The standard being 100 gms., the average threshold was at the Elementary School 8·75 gms. (M. V. 1·5, extremes 6 and 16 gms.); at the Preparatory School 9·3 gms. (M. V. 1·6, extremes 5 and 11·5 gms.)—the three sections of the former group averaging 8·5 gms. (clever boys), 9·5 gms. (average boys), 8·1 gms. (infra-normal boys); and the imbecile boy 4 gms.

Hence, connexion between Weight-Discrimination and Intelligence seems to be either zero or even slightly inverse; and boys of superior cultural status are hardly as acute in distinguishing fine differences of weight as those of lower social status.

The latter conclusion is in harmony with the results of anthropological investigation. Mr McDougall found the power of discrimination of small differences of weight rather more delicate in the Murray

Islanders than in Englishmen, the average least perceptible difference among the former being 3·2 °/₀ of the total weight, among the latter 3·9 °/₀[1]. But the latter conclusion seems at first sight inconsistent with the results obtained by Dr Spearman. At a village school the correlations between weight-discrimination of the 24 eldest children individually tested and two gradings for "Common Sense" and one for "School Cleverness" were (raw) ·38, ·27, and ·38; (corrected) ·43 (prob. error, about ·10). In a high class Preparatory School the correlation between a collective test of weight-discrimination and amalgamated school place, modified to eliminated age, was ·12 (prob. error ·09). Calculated from his figures, given in terms of 200ths of the standard, the average thresholds prove to have been, at the former, 11 (M. V. 4·4, extremes 4 and 28); at the latter, 12 (M. V. 4·1, extremes 4 and 28),—the apparatus used having been cartridge weights of 1000 grains and upwards, the heavier weights successively increasing in geometrical proportion[2]. It seems evident that the threshold determined in the two different investigations was not the same kind of threshold. The main differences in the method of determination (besides the fact that the subjects required differences of weight four times as fine as ours, *loc. cit.* p. 246) seem to have been that, in the case of his Preparatory School, the subjects were examined collectively, so that "it was impossible to control whether they all handled the apparatus in the same manner"; and, in the case of his village school, "the beginning of the test was devoted...to quietly affording the reagent a maximum of fore-exercise," and "before taking down each reply, a chance of reconsideration was given by repeating the test in such a manner as to reverse the constant error of time and space." The comparative lowness of the thresholds recorded by Dr Spearman is thus explained by the fact that they were obtained by "only recording those answers which were given under the most favourable conditions[3]," whereas our thresholds were obtained under conditions which were unfavourable to the objective truth of the judgments because they aimed at excluding every ground for the judgments except subjective sensations obtained in a manner as nearly as possible the same for all. Whether rigidity and limitation of condition is, more nearly than freedom and favourableness, 'the same for all' subjectively

[1] *Loc. cit. sup.* p. 198. Dr Spearman's village children were also slightly more acute at weight-discrimination than his preparatory children.

[2] *Amer. Journ. Psych.* Vol. xv. p. 286 sq.

[3] *Loc. cit.* pp. 249, 247; the last quotation in its context refers only to the group tested individually.

as well as externally is hard to decide. In a prolonged investigation into absolute thresholds, variety of conditions as to 'time' and 'space' would be indispensable; but in a comparative test lasting but 15 minutes differences of thresholds obtained with opportunities unrestricted in these respects probably indicate differences in power of inference or in rapidity of practice as often as differences of mere sensory acuity. In this way it seems possible to explain the apparent divergence between our respective results with this test[1].

(3) Sound Discrimination.

Apparatus and Procedure.

This test was introduced as an after-thought, somewhat late, into the programme. It was in consequence impossible either to repeat the observations on this test or to work through a preliminary set of experiments with a view to determining the most satisfactory instrument and the most satisfactory procedure for the purpose. It seemed that at least the results would be more comparable with those of the most recent and most successful employers of this test, if the form of the instrument and the form of the problem employed by them were adopted for the present series.

The tones, therefore, were produced by plucking successively the two wires of the dichord devised by Spearman, and the subjects were required to state whether the second note was "higher" or "lower" than the first.

The constant wire was tuned to E above middle C ($e' = 320$ vibs.); the variable wire was tuned to the same pitch by adjusting its tension till no beats were perceptible between the note emitted by it and the note emitted by the constant wire. Its length was then varied by moving the sliding clamp. This clamp is furnished with a vernier, moving upon a millimetre scale, so that the length of the wire vibrating

[1] Commenting on this passage Dr Spearman has written to me: "It seems to me that your results are in perfect harmony with that obtained at my high class preparatory school, as the correlation there was less than twice the size of the probable error, and therefore without significance. But at my village school, unlike you, I tried to include as much practice as possible, and this I think amply explains our at first sight discordant results; I have found over and over again that practice may greatly increase tendencies to correlate." By 'practice' is here apparently meant practice during *the same* sitting, or 'fore-exercise.' Practice in *successive* sittings seems to have the reverse effect on intelligence correlations in tests whose procedure is readily grasped. Cf. p. 168.

can be adjusted correctly to 0·1 mm. In the region of the once-accented octave a difference of 1·0 mm. corresponds to difference of about one vibration. The procedure was again the combination of Right and Wrong Cases with Minimal Changes,—commencing from an easily noticeable supra-liminal difference of about 10 vibration differences, and descending by steps of 1 vibration difference,—and 'without knowledge' at the critical stages.

Results.

One series was worked through at each school by the writer; and at the Elementary School 18 of the boys were also tested by Mr Flügel. For the two series thus obtained from these 18 the correlation coefficient is ·67; this figure is accordingly used as the measure of reliability both at the Elementary and at the Preparatory School. The reliability coefficient for the differential thresholds for pitch obtained by Krueger and Spearman from their eleven adult German subjects was much higher, viz. ·87[1]. The low 'reliability' of the results obtained at the schools seems due to two main causes. Though admirably adapted to collective work, or individual work with small groups, for which it had previously been employed, the form of dichord used by us was not so well adapted to the individual testing of groups so large that the experiments lasted over periods of several hours on several successive days, since the tension of the wires, and consequently the pitch of the notes, was found to vary appreciably from hour to hour and from day to day[2]. It was found, too, that successful judgments of *direction* of differences of pitch are not only intrinsically more difficult than judgments of mere difference, but also presuppose a knowledge of, and familiarity with, the meaning of the terms 'higher' and 'lower' as applied to the musical scale,—a knowledge which is not always available in children who have not enjoyed some degree of musical training. Owing to the late introduc-

[1] *Zeitschrift für Psychologie*, Bd. 44, p. 77.

[2] On finishing the experimental work recorded in the present section we were inclined to believe that tuning-forks would have given more reliable results; but subsequent experience with these convinces me that the rise in pitch of the tuning-fork after being sounded, unavoidable unless complicated and expensive accessory apparatus is used, renders its tones even more unsatisfactory for work such as the present than the ordinary dichord, while the manipulation of the dichord is easier to acquire. In his report upon "Apparatus and Methods for the Experimental Investigation of Sensations of Tone," for the American Psychological Association, Dr Spearman proposes several modifications of the original form of his instrument which will minimise or abolish these and other slight imperfections that he notes.

tion of this test into our programme we were unfortunately unable not only to repeat the series for the mathematical elimination of errors otherwise unavoided, but also to make preliminary experiments which would have led to an earlier discovery of avoidable sources of error, such as the two just mentioned.

The correlation between the thresholds for pitch (amalgamated where possible) and the Headmasters' order is, for the Elementary School ·40 (p.e. ·10), and for the Preparatory School ·37 (p.e. ·17). The raw correlations average ·40 and ·35. These being distinctly more than twice the size of the Probable Error, we may legitimately apply the correction formula, when we obtain pure coefficients of ·52 and ·41. The correlations obtained by Krueger and Spearman between Pitch-discrimination and Adding, the *Kombinations-Methode,* and Memory were (av. raw) ·67, ·59, ·17; (corrected) ·68, ·64, and ·00 respectively. At Spearman's village school the correlations of Pitch-discrimination with "Common Sense (A) and (B)" and "School Cleverness" were ·44, ·41, and ·25, yielding a 'pure' correlation with General Intelligence amounting to ·71. At the high class Preparatory School the correlation of Pitch-discrimination with Average School place was ·33, and with proficiency in four branches of study ·51 (av. raw); omitting non-musicians and 'correcting' ·87. These results harmonize with our own. The higher coefficients may be due either as before to a slight difference of psycho-physical method affecting the practice or procedure of the subjects, or perhaps to a more successful manipulation of the instrument[1].

The average threshold at the Elementary School was 6 v. d. (M. V. 1·9 v. d.; extremes 1 and 11·5 v. d.). The average thresholds of the three sections were 4·6, 5·6, and 7·1 v. d. The threshold of the imbecile was not, on the most generous estimate, lower than 10·0 v. d.; he was, however, peculiarly difficult to test. One or two boys who were known as "hard of hearing," that is, unable to hear sounds of low intensity, proved to have a high threshold also for pitch; they usually ranked as somewhat dull scholars. The average threshold at the Preparatory School was 3·5 v. d. (M. V. 2·2 v. d.; extremes 0·3 v. d. and 8·0 v. d.). The average thresholds of Krueger and Spearman's 11 German adults was 7·5 v. d. (M. V. 2·2; extremes 1·4 and 25 v. d.), of Spearman's 24 village children 9·8 v. d. (M. V. 6·0; extremes 1·3 and 30 v. d.), of his Preparatory boys 4·6 v. d. (M. V. 3·0;

[1] Dr Spearman writes: "Myself, I should attribute it to the same cause as before, viz. more fore-practice."

extremes 0·6 and 20 v. d.)[1]. As to hearing among savages, the available data "so far as they go point to inferiority rather than superiority in Papuans and Todas[2]." The average thresholds for pitch-discrimination among the Papuans tested by Dr Myers were 15·4 v. d. (adults, first sittings), and 12·5 v. d. (children, second sittings); those of the inhabitants of Aberdeenshire tested by him were 7·6 v. d. (adults, first sittings) and 4·7 (children, second sittings)[3]. The Todas were tested by Dr Rivers only for auditory acuity, not for pitch-discrimination. The present results, therefore, harmonize with those of previous investigators in finding an acuter judgment of pitch among groups of superior cultural status. This difference was not due to superior practice on the part of the Preparatory boys, since in this respect the Elementary boys were equally, if not more, fortunate. All of them enjoyed musical instruction; more than half were choristers; many learnt some musical instrument; whereas five out of the thirteen Preparatory boys neither sang nor learnt music, and had to be told the meaning of the terms 'higher' and 'lower.' Beyond the fact that this unfamiliarity with the nature and nomenclature of their task seemed in several cases unfavourably to handicap them, lack of practice seemed to make little or no difference. The average threshold of the boys who neither sang nor played was 3·4 v. d., that of those who played 3·5 v. d. (At Spearman's Preparatory School, however, the median threshold of non-musicians was 5·0 v. d., that of musicians 2·3 v. d.) This harmonizes with Dr Spearman's opinion that superior auditory acuteness is due to 'general culture' more than to special practice[4].

(4) Comparison of Lines.

Apparatus and Procedure.

The only test essentially involving vision was a test of the discrimination of the lengths of lines by an active method of comparison.

The apparatus used was that devised by Dr Rivers[5]. It consists of an oblong board, of such a size and weight that it can be conveniently

[1] Calculated from figures given, *Zeitschr. f. Psychol. l.c.* p. 72, and *Amer. J. Psychol. l.c.* pp. 286, 290.

[2] W. H. R. Rivers, "Observations on the Senses of the Todas," *Brit. J. Psychol.* 1905, Vol. I. Part 4, p. 391.

[3] *Cambridge Anthropological Expedition, l.c.* p. 168.

[4] *Amer. J. Psychol. l.c.* pp. 29, 30.

[5] W. H. R. Rivers, "Observations on the Senses of the Todas," *Brit. J. Psychol.* 1905, Vol. I. Part 4, p. 349.

held with the left hand, perforated with two pairs of holes, all four in the same horizontal straight line. Through either pair of holes runs a loop of string kept taut by a spring at the back. Half the length of each of the two strings is coloured white, and the rest black. They thus present the appearance of two straight horizontal white lines on a black surface, whose length can be varied from 0 to 20 cm. by moving the loops. In the following experiments the left-hand line was kept at a constant length, viz. 10 cm., to serve as the standard; and the subject was required to adjust the right-hand string, holding the board in his left hand and adjusting the variable line by moving the loop from behind the board with his right, and thus diminishing it from a greater length, or increasing it from a less, till he could no longer distinguish any difference of length between the two lines. The length of the lines formed by the white portions of the strings was then measured by a T-square furnished with a millimeter scale.

The procedure was 'without knowledge.' Each boy gave two sittings to this test, one to each operator. Ten crude observational figures were obtained from each boy on each occasion, five by the shortening or 'descending' procedure, five by the lengthening or 'ascending' procedure. The average, taken regardless of sign, of the differences between these ten figures and the standard was taken as the measurement of the boy's Average Crude Error. The boy's average determination was calculated from the ten figures taken together and the difference between this average, and the length of the standard was taken as the measurement of his Constant Error. The arithmetical mean of the deviations about this average was taken as the measurement of his Mean Variable Error.

Results.

The correlations with Intelligence and the group-averages for these three measurements are given in Table II. As the differences between the two operators' results are striking, they have been calculated separately as well as in amalgamation. The Constant Error may be either positive or negative; in using it as a basis for further calculations, determining its average for a given individual, or comparing its measure as obtained by two different operators, the magnitude of the error may alone be considered, or its direction may also be taken into account, reckoning negative amounts as amounts below zero. Both methods have been used in computing amalgamation and reliability coefficients.

TABLE II. *Comparing Lines.*

Averages of ten.

	Elementary School	Preparatory School
(1) Constant Error:		
(*a*) Reliability coefficient		
(i) Regarding only size of error	·34	– ·05
(ii) Regarding both size and sign	·22	·07
(*b*) Correlation between F's series and Headmaster's order	·14	– ·19
(*c*) ,, ,, B's ,, ,, ,,	·00	·51
(*d*) Correlation of amalgamated series and Headmaster's order		
(i) Individual averages calculated regarding only size of error	·14	·28
(ii) Individual averages calculated regarding both size and sign	·14	·05
Average size of error for group (F's series)	·26 cm.	·28 cm.
(B's series)	·32 cm.	·18 cm.
(amalgamated)	·29 cm.	·28 cm.
(2) Average Crude Error:		
(*a*) Reliability coefficient	·50	·54
(*b*) Correlation between F's series and intelligence (Headmaster's order)	·28	– ·07
(*c*) Correlation between B's series and intelligence ...	·17	·54
(*d*) Correlation of amalgamated series and intelligence ...	·29	·17
Average for group (F's series)	·41 cm.	·50 cm.
(B's series)	·40 cm.	·28 cm.
(amalgamated)	·41 cm.	·39 cm.
(3) Mean Variable Error:		
(*a*) Reliability coefficient	·23	·56
(*b*) Correlation between F's series and intelligence ...	·31	·01
(*c*) ,, ,, B's ,, ,, ...	·13	·56
(*d*) Correlation of amalgamated series and intelligence ...	·31	·35
Average for group (F's series)	·33 cm.	·40 cm.
(B's series)	·26 cm.	·23 cm.
(amalgamated)	·30 cm.	·31 cm.

The correlation between this test and Intelligence is highest in the case of the Mean Variable Error, and lowest in the case of the Constant Error. But in both these measurements the results of the two operators differ considerably, and the difference is itself different in the case of the two different schools. In the case of the Average Crude Error—the divergence, especially at the Elementary School, is not so great; the reliability coefficients are high, and about equally high, at both schools. The anomalies thus seem to be brought out by the further calculations; and consequently the Average Crude Error is taken throughout as the least precarious measurement of proficiency at this test for the purposes of the present investigation.

The correlations of the amalgamated A. C. E's. with Intelligence is ·29 at the Elementary School and ·17 at the Preparatory School. The average raw correlation between the two gradings for the A. C. E. and the various gradings for Intelligence proves to be somewhat higher, namely, ·34 at the Elementary School and ·31 at the Preparatory School, correcting to ·51 and ·44 respectively. As to the comparative merits of the two schools in the two tests, the amalgamated A. C. E. show a very slight difference in favour of the Preparatory School; the group averages being ·41 (M. V. 107) and ·39 (M. V. 179); but if we take the more satisfactory of the unamalgamated series for each school, then the difference is decided. So far as I am aware, no measurements have been made by means of this apparatus among savages. In estimating length of lines by eye with other methods, the Murray Islanders seem to have been slightly inferior to English students and children. The visual test employed by Dr Spearman was discrimination of luminosity. Figures obtained by previous workers, therefore, are here not available for purposes of comparison.

The origin of the curious divergences between the figures for the two operators and the two different schools is a matter for speculation. Such speculations, however, would be relevant rather to a discussion of the methodology and psychology of the test, than to a discussion of its relations with Intelligence. I append, therefore, a further table of

TABLE III. *Comparing Lines.*

Averages of five.

	Number tested	Average			M. V.			No. of average determinations	
		Flügel	Burt	Amal.	Flügel	Burt	Amal.	above 10	below 10
Shortening.									
Elementary boys ...	30	10·11	10·06	10·08	·276	·211	·225	37	23
Preparatory boys ...	13	10·23	10·07	10·15	·250	·203	·226	16	10
Cambridge adults ...	20	(Rivers) 10·35	—	—	(Rivers) ·168	—	—	16?	4?
Lengthening.									
Elementary boys ...	30	9·95	9·91	9·93	·235	·216	·227	25	35
Preparatory boys ...	13	9·80	9·83	9·82	·204	·180	·192	13	13
Cambridge adults ...	20	(Rivers) 10·12	—	—	(Rivers) ·144	—	—	16?	4?

calculations, showing the differences between the Averages and Mean Variations when calculated, not for the ten determinations together, but for the five determinations by the shortening and lengthening procedures respectively, and leave inferences to the reader. I have printed Dr Rivers' own figures as obtained with this test by a similar method upon twenty English adult observers (graduates and undergraduates of Cambridge)[1].

It may be mentioned that there was certainly a difference in the several experimental series as to the nature of the instructions to the reagents and the way in which they were carried out. In Dr Rivers' experiments the observer was apparently asked to shorten or lengthen the variable line "till it appeared to him to be of the same length [as the standard line]. Fine adjustment was allowed when the lines were judged to be approximately equal" (cf. *l.c.* p. 340). The boys on the other hand were told gradually to shorten or lengthen the variable line till they could see no difference. To forbid fine adjustment, however, was with the subjects and apparatus in question, rather a counsel of perfection: and the strictness of the two superintendents, and the strictness at the two schools, was probably not the same. These differences in procedure are probably the key to instructive divergences between the results of Dr Rivers and our own, between the results of Mr Flügel and myself, and between the results of the two schools. It will be noted that the under-estimation of the variable line when lengthened is on the whole not so great as the over-estimation of the variable line when shortened. This seems to indicate that the tendency to over-estimate the variable line, noted by Dr Rivers as operating with both ascending and descending methods, still operated with both methods in our experiments, though partially obscured by the constant error of stopping too soon. "The explanation of this general tendency," says Dr Rivers, "is doubtful." I should venture to suggest that it may be due to the fact that the centre of the board is not always held in the direct line of sight. There is a tendency to hold the variable line—left or right horizontal, according to the nature of the experiment—opposite the eyes; in consequence of this the standard line is seen somewhat foreshortened. Doubtless the larger discrepancies in the difference between the constant errors as obtained by the two observers at the two different schools are due to variations in the strictness with which the boys were made to hold the board exactly in front of the face. This general lack of pre-arranged uniformity in the procedure of the test is probably

[1] *Brit. Journ. Psychol. l.c.* p. 350.

the reason of the occasional low reliability coefficients, rather than any serious unreliability inherent in the nature of the apparatus or test itself. But the reliability at its best is in this test nowhere very high.

Conclusions (Simple Sensory Tests).

On comparing the correlation coefficients, we have the following generalisations with regard to the relations between Sensory Discrimination in various departments, and Intelligence. There appears to be no general connexion between Intelligence and capacity to discriminate Weights; any general connexion between Intelligence and Tactile Discrimination, if it exist, is of the slightest; there is considerable general connexion between Intelligence and Pitch Discrimination; and an undoubted general connexion between Intelligence and Visual Discrimination of lengths, though not to such an extent as in the case of Pitch Discrimination. Owing, I am inclined to suggest, to some participation of intelligence in his subjects' general attitude and procedure, as well as in the function specifically tested, the correlations obtained with school-children by Dr Spearman were larger than those obtained in the present investigation; nevertheless, they follow the same order, namely, Weight, Vision, and Sound[1].

It is to be observed that absence of general connexion between Sensory Discrimination and Intelligence in any particular department of sense is compatible with, and not refuted by, cases in which the general fineness of nervous organization, manifesting itself at higher levels as Intelligence and as superiority in more complex tests, is also shared by lower levels of psycho-physical constitution. Thus the most brilliant of the Preparatory group, a son of a mathematician and scientist of international repute, and therefore presumably inheriting the high order of intelligence his work at school displayed, was third in the Lines test, second in the Skin test, and an exceptional first in both Weight- and Pitch-discrimination; and these performances seemed to the onlooker to be due rather to peculiarity of general sensory organization than to peculiarity of attention or other feature of general mental attitude. The superiority of the congenital imbecile could certainly not be attributed to any such central factor. Accordingly, the correlation found between Intelligence and Pitch Discrimination and—to a less extent—

[1] The correspondence is not due to suggestion, as at the time of our actual experiments I had not noticed the order obtained by Dr Spearman, and before actually calculating the coefficients believed we were finding no correlation throughout the sensory region.

between Intelligence and Visual Discrimination both by Dr Spearman and by the present investigators are probably to be explained not as manifestations of a fundamental identity between Intelligence and General Sensory Discrimination, but rather historically, namely, by the large dependence of the development of intelligence in mammals upon visual acuity, especially in relation to the perception of space, and by the yet larger dependence of the development of intelligence in man upon power of speech, and of this in turn upon auditory acuity.

General Intelligence, then, shows little or no relation to senses which to civilised man are of low cognitive value; but it shows a marked relation to those senses which aid the perception of relations or formation of concepts, and are of high cognitive value.

Motor Tests.

Motor ability has commonly been tested by means of Reaction-Time apparatus—apparatus too costly and elaborate for extensive use at schools and unsuited to testing the easily distracted minds of boys. Among the motor tests not requiring unusual apparatus, those claiming to yield the most successful results are the Tapping test introduced by Binet, and the Card tests introduced by Jastrow.

(5) Tapping.

Apparatus and Procedure.

Binet's test of "Petits Points," which has been adopted also for this purpose by Bagley and Kirkpatrick, consisted of tapping rows of dots with a pen or pencil, thus recording the number of taps executed in a given time. I found, however, an undesirable amount of variation in the way in which the boys used the pen or pencil—some being accustomed to hold them in a way that rendered successful 'tapping' almost impossible, others remedying the tendency of these instruments to leave indistinct marks by moistening the point of the pencil or scratching in the dot with a circular movement of the pen. For the boys whose work is recorded here I therefore substituted a blunt needle mounted on a holder. They were instructed to hold the needle almost vertical, and to prick at maximum speed a line of holes in a sheet of paper laid over a thick cloth. Each was allowed a preliminary practice-

row; the number of perforations pierced in 15 secs. was taken as the measure of his rapidity.

Two experimental series were carried out at the Elementary School and one at the Preparatory School.

Results.

The reliability of this test was not high; the correlation between the two series, the operator being the same for each, was ·51.

The correlation of the average rate of Tapping with Intelligence was ·47 (p. e. ·09) at the Elementary School, and the corresponding correlation at the Preparatory School ·41 (p. e. ·16). The average of the raw correlations between the gradings for tapping and for intelligence were ·44 and ·28 respectively; correcting to ·65 and (assuming at the Preparatory School the same reliability coefficient as at the Elementary School) ·41.

The average speed of the Elementary group was 80·5 taps in 15 secs. (M. V. 8; extremes 59 and 97 secs.), of the Preparatory group 95 secs. (M. V. 11; extremes 66 and 119 secs.). The correlation of Tapping with Athletics according to rough gradings by the masters was ·35 at the former school and ·14 at the latter; but its correspondence with athletic constitution, as distinguished from athletic interests or achievements, and estimated roughly by eye and by the information contributed by the masters, seemed much closer, particularly at the Preparatory School.

The reliability coefficient is perhaps lower, and the correlation with intelligence perhaps higher, than might have been anticipated. Subsequent experiments in 'tapping' carried out by means of a rotating cylinder, which carries the padded paper past at a uniform speed so that the only movement the subject has to make is the movement of pricking itself, have yielded distinctly higher reliability coefficients. Apparatus was not devised or used for the experiments upon the school children for tests in which it could be dispensed with. Consequently in the Tapping test the child besides pricking had also to move his arm along the paper, and sometimes back again, *βουστρόφηδον*. In this way there was introduced into the movements the need of more accurate co-ordination. And to this both the slight peculiarities in the results seem attributable.

(6) Dealing.

Apparatus and Procedure.

The apparatus used consisted of playing-cards[1] of the kind made by De La Rue and Co., called 'Pneumatic Cards,' the backs being roughened by grooves, which allow air-spaces between the cards when the pack is held in the hand, and thus facilitate dealing and prevent misdeals. The subject was required to deal 50 cards into five heaps in the ordinary manner, and the time was recorded by a stop-watch.

Three series were made with the larger group, and two with the smaller—Mr Flügel again undertaking one at each school.

Results.

The reliability coefficients of these series were high, namely ·88 and ·80. The coefficients of the amalgamated series when correlated with Intelligence were ·44 (p. e. ·10) and ·29 (p. e. ·18); average raw correlations ·48 and ·05; corrected coefficients ·54 and ·06.

The average time taken to deal the pack of cards was at the Elementary School 31·6 secs. (M. V. 5·9; extremes 20·3 and 46 secs.); the averages of the three sections of the group were,—clever boys 26, average boys 31·6, dull boys 35·6, the time taken by the imbecile boy 46, secs.,—thus confirming the presence of a marked correlation. The average time at the Preparatory School was 27·3 secs., or nearly as fast as the 'clever' section of the Elementary group (M. V. 1·8; extremes 20 and 34 secs.).

On plotting the 'surface of distribution' of the individual figures for the two groups, the curve limiting the surface of distribution for the Elementary group was found to be fairly symmetrical, and both 'median' and 'mode' (most frequently recurring speed) coincided with the average; while the curve for the Preparatory group was found to be decidedly asymmetrical or 'skew,' and the median (26·5 secs.) and mode (26 secs.) both distinctly below the average, there being but one figure still smaller than the mode. A small number such as that of the Preparatory School group naturally tends to evince irregularities and asymmetry in the curves constructed for that group. But a difference in symmetry so marked, together with the difference between the respec-

[1] The method of studying reaction-times by playing-cards was introduced by Jastrow (*Science*, Vol. VIII. p. 237), and has been successfully used by Bergström (*Am. J. Psych.*, v. p. 356). For an application of these two forms cf. also W. G. Smith, this *Journal*, Vol. I. Part 3, p. 244. Their application to testing intelligence is, I believe, new.

tive correlations with Intelligence, seems to indicate the presence of some extraneous factor not appreciably influencing the Elementary boys, but influencing the other group so as almost to antagonise any tendency to correlate with Intelligence. Such a factor naturally suggested itself in previous practice. It was conceivable that in the main only the more intelligent subjects at the Elementary School habitually played cards and dealt for themselves, and that only the less intelligent of the Preparatory boys spent their time in such indoor recreation. Enquiry was accordingly made of the boys as to the amount of their familiarity with cards, and they were graded in rank on the basis of their replies. This rough grading correlates with the experimental results to the extent of ·44 in the Preparatory School, and of only ·28 in the Elementary School. Here then was the cause of the difference in distribution frequency of the observational data, and of the correlational coefficients; its operation would be explained if at the lower class school the brighter boys and at the higher class school the duller boys were the chief regular card-players. There proved, however, to be no reliable positive correlation between Intelligence and familiarity with cards at the Elementary School, nor any negative correlation at the Preparatory School. The difference of incidence therefore operated in some other way than that expected. It transpired that at the Elementary School the card-players played at home "in the winter evenings," "during the holidays," and thus did not play regularly and did not necessarily deal for themselves. At the Preparatory School the more frequent card-players were boarders, who played at school "on most wet afternoons" and necessarily dealt for themselves. Thus at the latter school practice had simply had a far more profound influence, and (if we disregard the possibility of some abnormal disturbance in this test at this school—and of this there were no signs) this profounder practice probably determined the results of the experiment so completely as to overlay and obscure the original connexion between Motor Rapidity and Intelligence.

Conclusions (Simple Motor Tests).

Of the two forms of simple motor test, Tapping seems a more satisfactory method than Dealing, especially as its defects could largely be remedied by improved apparatus, while those of Dealing cannot. Motor tests seem to have a higher correlation with Intelligence than

Sensory tests. But where motor rapidity is due to frequent practice[1], as in the Dealing of the Preparatory group, the correlations with Intelligence and other tests are reduced, abolished, or inverted. Thus so far as motor rapidity is the function of temporary 'facilitation' of the paths of neural discharge it appears also to be a function of intelligence, while so far as it is a function of permanent 'canalisation' of those paths it but slightly or inversely related to intelligence. Facilitation, however, is a function of operative attention; while canalisation, though due to the operation of attention in the past, corresponds with diminution or absence of attention, as the adaptations of the past become the habits of the future. Here, then, the correlation between the tests and Intelligence seems more direct, and more likely to be the outcome of a central factor, than in the case of the sensory tests.

Sensori-motor Tests.

In the two following tests the reaction was complicated by sense-perception, or by recognition.

(7) Card Sorting.

Apparatus and Procedure.

Fifty pneumatic playing-cards, consisting of ten from each of five different packs with backs of five different colours, were arranged in a haphazard order, special care being taken to avoid favourable or repeated sequences. The subject was then required again to deal them out into five heaps against time, now distributing them according to the five colours. The positions and colours of the heaps thus to be formed were indicated for him by five similar cards placed conveniently in a single row on the table. In Sorting and in Dealing the boy was instructed not to correct mistakes in allocation or misdeals; if made, they were counted as so many seconds to be added to his time, according to his average speed per card.

Three experimental series were undertaken at the Elementary School and two at the Preparatory School as before.

[1] The correlations for the two successive series of experiments in Tapping (given in Table VII) are at first sight not consistent with this inference. But the effects of a single attempt at 15 seconds Tapping scarcely constitutes practice in the sense implied above; and the general effect of repetition not only with Dealing, but with most of the other tests, seems to confirm what is there said.

Results.

The reliability coefficients were ·84 and ·38; the correlations with Intelligence ·52 (p.e. ·09) and ·56 (p.e. ·13) for the amalgamated series, ·45 and ·60 raw, and ·53 and 1·63 corrected. A correlation larger than unity is, of course, impossible. The last coefficient, therefore, contains an error of at least ·63,—an error nearly three times as large as its 'probable error,' which works out at ·23[1]. This, however, proves not so much that 'correction' is invalid, as that the formulae are at present sometimes inadequate. The low reliability coefficient and high intelligence correlation which, as compared with those of the Elementary School, characterise the Preparatory, are possibly due to the fact that the relative novelty involved in the process of sorting was for them greater, since the process of dealing was to them more habitual than at the other school. The correlations between Sorting and Dealing (amalgamated series) is at the Elementary School ·72; at the Preparatory School it is only ·02. The correlation between Sorting and previous practice in handling playing-cards proves to be even slightly inverse, namely,—·11 (Elementary School)[2]. Sorting for the habitual dealer doubtless involves not merely the acquisition of the complex tendency to lay a given card upon the heap of similar colour, but also the inhibition of the simpler and older tendency to lay it on the heap next in order after the heap on which the preceding was laid. Hence, among the more practised card-players of the Preparatory boys there was probably room for a greater play of at least one of the functions credited to 'intelligence,' and at the same time for a greater diversity and a greater irregularity of operation on the part of this additional influence. When projecting the entire series of tests it was hoped that it would be feasible to discount by subtraction the element of practice with cards from the results of the experiments in Sorting by means of the results of the experiments in Dealing. But it became evident in the course of the observations and calculations that the processes are involved in a way which is scarcely so simple as to warrant such a procedure; indeed we have already noted that experimental Dealing is no invariable measure of practice in handling cards.

[1] According to the formula for p.e. of corrected coefficients given on p. 111.

The six other corrected coefficients printed on pp. 176, 177 as 1·00, also actually work out to figures considerably more than unity.

[2] I find that Miss Thompson noted a similar fact in her experiments with Jastrow's Card-sorting upon students at Chicago: "Those who made the best records, both men and women, were people who played cards little or not at all." *Mental Traits of Sex*, p. 16.

At the Elementary School the average speed in this test was 48·1 secs. (M. V. 6·3). The fastest was 34 and the slowest 61·6 secs. The average speed for each of the three sections 41, 49·3, and 51·5 secs., and the speed of the imbecile child 87 secs. At the Preparatory School the average speed was 41·3 (M. V. 4·1)—again nearly as fast as that of the brightest section of the Elementary boys. The fastest was 36·5 and the lowest 49·5 secs. Thus whereas in Dealing the fastest speed achieved at the Preparatory School was slightly greater than that attained at the Elementary School, in Sorting it was considerably less; this seems to confirm the supposition that the complication of the afferent part of the reaction by the introduction of the element of Sorting as such entailed a greater interference with the reaction process for the Preparatory boys than for the Elementary boys.

(8) Alphabet Sorting.

Apparatus and Procedure.

This test is believed to be entirely new, and was devised by Mr McDougall. The apparatus required consists simply of a stop-watch, and two complete alphabets of childrens' cardboard letters, such as are sold for the game of 'Word-making and Word-taking.' Each letter of the alphabet is boldly printed upon a single white card about 20 mm. square. The fifty-two cards were laid before the subject upon a table, placed right way up, but arranged in an irregular or 'chance' order in three rough rows. The subject was not allowed to see them until he commenced the experiment; and was instructed to pick out in order, and arrange in sequence in two rows below, one complete alphabet from the two before him. He was directed to work through the alphabet in order continuously, not to pick out letters as his eye fell upon them, and was started by a pre-arranged signal. The task was, as before, to be performed at maximum speed; and the time occupied was recorded by the stop-watch.

Three series were executed at each school, one of each set being as before undertaken by Mr Flügel.

Results.

The average reliability coefficients, ·60 and ·48, are not quite so unequal as for sorting, but still both somewhat low. This is probably

due to the opportunity afforded by the nature of the test for slight differences in the operator's method of arranging the letters and carrying out the test, and for the operation of chance generally. Probably for the same reason the several raw correlations of the three successive series are somewhat widely discrepant (cf. p. 168).

The average raw correlations with Intelligence are ·50 and ·61 (correcting to ·68 and ·91), being thus at first sight much the same as those for Sorting. But the correlation relations between the amalgamated or average results show a far more marked improvement upon these than do those for Sorting; they rise to ·61 (p. e. ·08) and ·80 (p. e. ·07). The correlation with Intelligence (Headmaster's estimate) is found, too, to increase on each successive occasion on which the best was performed. Thus on the first occasion of the test (operator B.) the coefficient was ·45; on the second (operator F.) ·70; on the last (B. again) ·74. The separate reliability coefficients similarly improved. There was no such improvement observable in the case of Sorting. It would thus appear that the Alphabet experiment, involving as it does a similar type of reaction, tests much the same capacity as the Sorting experiment; and so far as it does differ from the latter, is on the one hand somewhat more closely connected with Intelligence than Sorting, particularly where repetition is possible to eliminate the element of chance, but on the other hand somewhat less reliable than Sorting, particularly where the differential influence of card-practice does not complicate the element of dealing. The Alphabet test also unfavourably handicaps those who have a difficulty in reading and therefore do not readily recognise letters; and the Sorting test unfavourably handicaps those who are more or less colourblind, and therefore do not readily recognise the colours on the backs of the cards; while both are unfavourable to those who are weak in the arm and often to the left-handed. Specimens of each defect will usually be present in groups as large as the larger of our groups.

The disadvantages of this test are thus not serious; similar disadvantages are inevitable in any test. The defects which seem to have lowered the reliability coefficients could largely be improved upon by devising stricter conditions for future applications. The test is a very quick and simple one, and so could easily be repeated often enough to yield results of considerable reliability, and therefore, even without correction, of very high correlation. Even in our own use of it, the results obtained were quite remarkable: on correction, it furnishes in both schools a higher average correlation with Intelligence and the

other tests, without any countervailing error of disastrous size, than any other test at the Elementary School, or any other test except Dotting at the Preparatory School[1].

At the Elementary School the average speed of the group was 91 secs. (M. V. 15·4, extremes 50 and 138 secs.), the three sections averaging 75·6, 90·5, and 103 secs., and the imbecile boy occupying a little over five minutes on one occasion, and six on another. At the Preparatory School the average speed was 74 secs. (M. V. 13, extremes 48·5 and 119 secs.), actually superior to that of the brightest boys of the Elementary School group.

Conclusions (Compound Sensori-Motor Tests).

Depending as they do for their performance upon processes of a more complex nature and a higher mental level, tests combining perception with motor reaction[2] seem to involve Intelligence to a still higher degree than relatively simple sensory or motor tests. Of the two above discussed, the Alphabet seems to be in practice far the more efficient.

At the outset of the entire investigation, the last three tests (the Dealing, Sorting and Alphabet tests) were applied to 63 boys varying from 10 to 16 years of age with a view to determining the nature of the correlations of the tests with Age. They prove to be ·14, ·45 and ·29 respectively, the probable error of a series of this size being about ·07. The peculiar nature of these correlations is probably to be explained by the character of the selective influences that had been at work among the boys at upper and lower ends of the scale of age. From these three samples it was concluded, firstly, that the correlations of our tests with Age would be sufficiently large to introduce an undesirable factor disturbing the estimation of their correlations with Intelligence; and consequently it was determined to eliminate this influence at the beginning of the work by a suitable choice of reagents; secondly, that the correlations were not large enough to vitiate the assumption that such tests as should prove to be correlated to a high degree with imputed Intelligence, such as the Alphabet test, might be regarded to a large (though not perhaps an equal) extent as correlated also with, and therefore tests of, *innate* ability; and, thirdly, that the only way of proving this assumption empirically would be to take the *same*

[1] See Table IX at end (pp. 176, 177), and Tables V and VI (pp. 161, 162).

[2] It perhaps may be doubted whether the Dotting test yet to be described should not also be classed under this head, since in the general nature of the process and in results it may be considered to resemble the Alphabet test.

reagents at a later period of their educational career and test them again with the same tests.

ASSOCIATION TESTS.

(9) Immediate Memory.

Apparatus and Procedure.

The Memory of the children was investigated according to the method employed by Prof. Meumann[1], who investigated the immediate memory of some 800 Zurich school-children by determining by a series of mass-experiments the maximum number of words which each child could reproduce after once hearing them read in class by the class-teacher in groups of 4, 5, 6, 7 and 8 words at a time. Two different sorts of words were used: namely, (1) words of *concrete* significance chosen from the children's circle of ideas, *e.g. Papier, Strasse, Ofen, Feder*, &c.; (2) words of *abstract* significance, less known to the younger children, such as *Menschheit, Gesetz, Masse, Organ, Anziehung*, &c. As a result of his experiments, confirming the casual observations which had suggested them, he concludes that the degree of mental development of the children is determined by their capacity of abstraction, and can be recognised by their retention of abstract words, since such retention is essentially dependent upon their power of understanding the words.

In applying this method to our purposes, the procedure used by Meumann was modified in the following respects. The children were, as always, tested, not by a class experiment, but privately and individually; and the test was conducted, not by the master, but by ourselves personally. The words chosen were all of one syllable, and, in addition to concrete and abstract words, a third series consisting of 'nonsense' syllables was used. The children not only heard the words, but also saw and spoke them—thus in some degree obviating differences of Imagery-type. In presenting the words to be read by the subject, no use was made of the machines usually employed in memory experiments, since the distraction caused by such pieces of apparatus, especially when unfamiliar to the child, more than counterbalances the accuracy gained by mechanically regulating the tempo, &c.[2] The

[1] E. Meumann, "Intelligenzprüfungen an Kindern der Volksschule," *Die Experimentelle Pädagogik*, 1 Band. Heft 1/2, 1905.

[2] This was found to be the case by Messrs W. McDougall and A. M. Hocart in some unpublished investigations on memory undertaken at the same school in 1905 with Müller's rotating cylinder.

words were printed by hand in 1 cm. Roman Capitals upon large cards in five vertical columns of 4, 5, 6, 7 and 8 words respectively, one card containing 30 abstract words, another 30 concrete words, and a third 30 nonsense syllables. In explaining to the boy beforehand the way in which he was to memorize and reproduce the words, he was warned to be careful to observe as far as possible in his reproduction the order in which the words were shown; and before each column he was told the nature and number of the words it would be found to contain. Over the card to be learnt was laid another and larger cardboard sheet, having an oblong aperture near its centre of such a size as to expose one and only one word at a time. By sliding this screen at a favourable rate over the printed card towards the boy, the words of each column were successively revealed through the opening and thus presented one by one to the view of the boy, who read them aloud as he saw them, and simultaneously heard them pronounced by the operator. As soon as a column had been presented in this way, the words contained in it were scribbled down by the boy from memory on paper previously ruled into appropriate columns. In this way 30 monosyllabic concrete nouns, 30 monosyllabic abstract nouns, and 30 meaningless monosyllables were seen, heard and read, 4, 5, 6, 7 or 8 at a time; and, as far as possible, reproduced by the boy in their original order, column by column.

The boys' papers were subsequently marked according to a system based upon Meumann's system of marking: each word correctly reproduced in its correct place counted 4, a correct word misplaced counted 3 if its position was altered by only one place, and 2 if removed by more than one degree; an incorrect word, if either initial consonant sound, or final consonant sound, or medial vowel sound alone was incorrectly altered, counted 3, if in its right place; 2, if one place removed, and so on; if two such components were altered, it counted 2 if rightly placed, and 1 if wrongly placed; the omission of a word, or the substitution of an extraneous word, counted 0[1].

Results.

The following are the figures for the three complete series made at the Elementary School, and for the two complete series made at the Preparatory School, based upon the total marks obtained at each sitting

[1] Meumann's full mark was 3, instead of 4, for each correct word correctly placed. To ensure that the method of marking should throughout be the same for all, the entire series was kindly marked by Mr Flügel.

by each individual for 'Concrete,' 'Abstract,' and 'Nonsense' Memory taken together:

TABLE IV. *Memory.*

	Elementary School	Preparatory School
Reliability coefficients	·70	·93
Correlation of amalgamated series:		
with Headmaster's order for intelligence	·57	·78
with Examination order (literary subjects) (mathematical subjects)	·67	·82 ·69
Average correlation of unamalgamated series with the various gradings for intelligence (Headmaster's)	·53	·80
(Assistant Masters')	·43	—
(Boys')	·46	·72
(Average)	·48	·76
Corrected correlation	·60	·82
Average no. of marks gained by boy per sitting	134·2	216
Mean Variation for group	27·7	29
Extremes (highest boy's average)	223·3	265·5
(lowest boy's average)	68·6	153·5
Average for the three sections: 'clever' boys	167	—
'average' boys	135	—
boys 'below normal'	134·2	—

From these figures it is clear, in the first place, that Immediate Memory is correlated to a considerable, but not a high, degree with Intelligence as estimated by the Headmaster's grading; to a slightly lower degree with Intelligence estimated by the general impression of others than the Headmaster; and to a significantly higher degree with Intelligence as estimated by the results of examination—particularly of the examination in 'literary' subjects (Classics, &c.). The correlation between the Headmaster's estimate of Intelligence and the examination order (on which, it is to be remembered, the Headmaster's estimate was with subsequent modification originally based) was, as we have seen, ·81 at the Elementary School, and ·78 (actually the same as the correlation with Memory) at the Preparatory School. The correlation between all other tests of intelligence and the results of the examination is much lower. Thus we seem to have scientific proof of what on *a priori* grounds has commonly been surmised, namely, that the present examination system tends to test mainly that aspect of intelligence which manifests itself in memory, to the neglect of other manifestations of intelligence, and to the inclusion of other factors of memory which distort even this manifestation of intelligence. There is clear evidence

also that at both schools the Headmaster's estimate is also biassed towards memory, since the correlations of the Memory test with the other provisional estimates of Intelligence are much lower.

This conclusion is corroborated when we attempt to trace the divergences between the figures for the two schools to their probable cause. The Elementary School which supplied our subjects, in common with many other elementary schools, had followed the reaction against excess of rote-work in its instruction to a degree which, in the opinion of the Headmaster, was even extreme; in its examinations, too, it aimed rather at setting problems, directly involving intelligence for their solution, than at demanding facts or formulae, primarily involving memory for their reproduction and intelligence only indirectly; and these examinations thus correlate more closely with the Headmaster's order and less closely with Memory than at the other school. A preparatory school, on the other hand, preparing boys for scholarship examinations at the great public schools, necessarily trains and disciplines to a high degree the memories of its scholars, particularly of its more intelligent scholars. This superior power of memorization on the part of the Preparatory School boys unmistakeably betrays itself in our experimental investigation. In no other test do the observational figures show such a complete superiority to those of the Elementary boys; the marks obtained even by the weakest boy in the former group was not much below the average marks obtained by the seven 'cleverest' boys of the larger group. The peculiarly high reliability coefficient measures the remarkable accuracy and steadiness of their work, which was also conspicuous during the actual experiment. For while the Elementary boys could not conceal their dislike for the Memory test, the Preparatory boys were here evidently most at home. In these boys—preparing as they were for examinations—susceptibility to the irrelevant factors liable to hinder subjects applying themselves to the assimilation and reproduction of a maximum amount of material in a limited amount of time had probably been largely reduced by previous discipline. Hence the results of their work would naturally correspond more closely with their Intelligence—particularly if, as seems possible, the efficient factor in Memory and Intelligence alike be attention manifesting itself in different guises. The fact that in this test attention appears to have been more efficient among the Preparatory boys than among the Elementary boys, seems to have made their reproduction of the associations between the words (as opposed to their reproduction of primary memory images of the words

themselves) more accurate[1]. The masters who drew up the grading for Intelligence at this school may have also been specially biassed in favour of an accurate and capacious memory as a symptom of Intelligence.

The coefficients of correlation between 'Concrete' Memory, 'Abstract' Memory, and 'Nonsense' Memory, taken separately with Intelligence, are, when the results of the successive sittings are amalgamated, ·58, ·48, ·43 respectively at the Elementary School, and ·84, ·78, ·75 respectively at the Preparatory School; thus the memory for abstract words does not show a higher, but a lower correlation, with Intelligence. The average marks were 60·7 (M. V. 11·4), 49·1 (M. V. 10·5), 25·2 (M. V. 7·5) respectively at the Elementary School, and 87·5 (M. V. 9), 76 (M. V. 11) and 52 (M. V. 12) at the Preparatory School; and thus the introduction of difficult vocables, whether abstract nouns or meaningless syllables, proves in both groups to be on the whole a distracting element. Indeed the only three cases where Abstract Memory gained better marks than Concrete Memory were those of boys, who so far from being boys of superior Intelligence, were placed 15th, 26th and 29th in the Headmaster's order. Hence so far as concerns children of the age and station examined by us, Meumann's claim that superiority of 'abstract' memory to 'concrete' memory is a strong mark of Intelligence is not confirmed[2].

(10) Mirror Test.

Apparatus and Procedure.

The 'Mirror' test is a new[3] test intended to measure the child's adaptability, i.e. his power speedily to acquire new co-ordinations of

[1] This point is of importance in considering the relations of the Memory test to the Sensory tests; fuller discussion of it is therefore postponed till the section dealing with the interrelations of all the tests (p. 167).

[2] In the children of the Volksschule investigated by Meumann, up to the age of 12 the actual errors made in reproducing the abstract words were in general greater than in reproducing the concrete words; with children of 14 and 15 years of age the relation reversed itself. *Exp. Päd. l.c.* p. 70.

[3] The suggestion of writing before a mirror is due to Mr Keatinge. For the present imperfect form of the test the writer is responsible. A similar experiment for a different purpose has since been published in Judd's *Laboratory Manual of Psychology*, Vol. II. p. 49. A prolonged investigation into adaptation to mirror-vision is described by G. M. Stratton, "The Spatial Harmony of Touch and Sight," *Mind*, Vol. VIII. p. 492, 1899. Mirror-drawing sometimes appears as a parlour pastime; and this may occasionally vitiate the test, as in the case of several of our Preparatory boys. One of the Elementary boys had also served as subject in a series of apparently analogous experiments carried out a couple

movement appropriate to circumstances relatively novel. The reagent traces over a geometrical pattern a number of times successively, both the pattern and his hand being seen by him, not directly, but in a mirror. The visual and motor factors at work in such a task involve in themselves nothing specially new or highly complex; but the relation between them is disturbed, and must be readjusted. The gradual improvement in the process of readjustment is indicated by a progressive reduction of the time required to make one complete tracing.

The apparatus employed was of the following nature. A sheet of paper was laid upon a thick cloth covering a firm table, and upon this paper was pinned a sheet of stout millboard pierced with eight holes arranged in the form of an octagon or circle 20 cm. in diameter, with a ninth hole in the centre. On piercing with a blunt mounted needle or 'seeker' through the holes in the card and through the paper beneath, a pattern will be left upon the paper similar to that upon the card. In this experiment, as in 'Tapping,' pricking holes with a seeker has several advantages over marking dots with a pen or pencil: it avoids accidents with the nib or pencil-point; it ensures that the method of holding and using the instrument shall be practically the same with all subjects; it leaves a comparatively infallible record of the subject's performance, and by simply renewing the paper a separate record for each subject can be obtained. Behind the pattern to be traced and facing the subject is fixed a mirror, not quite upright, but tilted towards the subject at an angle of 85° to the horizontal card; and the pattern is concealed from the subject's direct view by a horizontal screen, supported on a frame, underneath which his hand can freely move. Thus the holes, the needle, and the subject's hand can only be seen by reflection in the mirror. When about to commence the experiment, the subject holds the mounted needle in his right hand underneath the screen, and the point is inserted for him in the hole at the centre of the pattern. He is instructed to start at a given signal by piercing through this hole; then to find a certain hole on the circumference (marked with green), to pierce this; then to go to the hole next it (marked with red), and to pierce this; and so on in the same direction round the circumference of the circle till he reaches the green hole, when he is to return from this to the centre again. The time occupied in so doing is recorded by a stop-watch.

of years before at St Barnabas School, Oxford. According to his account, the test then consisted of pricking out letters of the alphabet as fast as possible; it was conducted by a stranger, whose identity and design I have been unable to discover.

At the same sitting, but with intervals of a few seconds, this operation is repeated till it has been achieved six distinct times. Such a six-fold series was obtained at both schools. Twelve weeks after the six-fold series, a second series of sittings was obtained at the Elementary School; in this series each boy's sitting included only two consecutive attempts.

Results.

In the six-fold series, six figures, recording the respective durations of the six successive tracings, were by the above procedure obtained for each individual. These figures provided the data from which to extract some comparable measure of his relative improvability so far as revealed in rapidity of adaptation during the entire sitting. A simple and satisfactory method of mathematically treating these raw figures, in order to obtain some such measure, is difficult to discover. Various procedures suggested themselves: to subtract the last and lowest figure from the first and highest; to subtract the average of the last three from the average of the first three; to express either of these remainders as a fraction of the original speed, or of the total time of the sitting; to reduce all sittings to terms of a common original speed, and calculate the coefficient of their curves; or simply to take the sum or average of the six experimental figures. Of these, the last method gives the highest correlation with Intelligence; while the others are in practice not so satisfactory, and are all more or less open to obvious objections on *a priori* grounds.

The correlation at the Elementary School between the total results of the first series of sittings and those of the second was ·52. From its very nature the Mirror experiment cannot strictly be held to test quite the same capacity on repetition. The observed figures indicate considerable retention by the subjects of the effects of the first sitting; and doubtless the degree of retention is not the same for all. Indeed the second series was undertaken partly with the hope that it might furnish a test of the retentiveness of improvability rather than of improvability itself. No method, however, presented itself of isolating this second capacity, and of differentiating the subjects accordingly. In any case the difference of retentiveness does not seem to have been great. And of the various substitutes for a reliability coefficient that one might devise, the figure cited above probably gives the best approximation for these particular experiments. In lieu of a better it is therefore employed at both schools for purposes of correction.

The correlation of results of the six-fold series with the Headmaster's estimate of Intelligence was ·67 (p. e. ·07) at the Elementary School, and ·54 (p. e. ·14) at the Preparatory School. The correlation of the two-fold series was ·22 (p. e. ·13). The average of the raw correlations with the various empirical estimates of Intelligence was ·50, and ·47, correcting to ·74, and ·68 respectively.

The average total time for the six-fold series was 388·9 secs. for the Elementary boys (M. V. 90·1; extremes 181 and 747 secs.; averages of the three sections 347, 383·7, and 429·8; time occupied by the imbecile boy 2464 secs.). The average for the Preparatory boys was 257 secs. (M. V. ·54; extremes 156 and 526 secs.). The divergence between the two schools is largely due to the fact that 4 out of the 13 Preparatory boys had had previous practice at an analogous task in the form of a not very common parlour pastime. Only one of the 30 Elementary boys had done any similar exercise before. The divergence might also be in part attributed to a greater familiarity with the use of the mirror among boys of the higher classes as compared with boys of a lower status. A similar factor apparently operates when the test is applied to children of the opposite sex, though subsequent applications to very young children, and to adults, have led me to wonder whether here we are not dealing with one of the uninvestigated innate differences between the two sexes[1]. The unexpected degree to which the effects of practice may persist was demonstrated in the later double series at the Elementary School. The average speed in this group was 103 secs. for the first tracing, 39·6 secs. for the sixth[2]; 34·5 secs. for the seventh tracing (i.e. after 12 weeks interval), and 27·4 for the last. The individual boys resumed the task with a speed in no case much slower, and in the case of 16 boys out of 26 actually faster, than that with which they had left off three months before. This furnishes a striking experimental parallel to the dictum quoted by Prof. James: that our brain learns to swim during the winter and to skate during the summer[3]. If the four practised boys are omitted from the Preparatory results the average speed is still somewhat greater, namely 291·8 secs., and the correlation with intelligence rises distinctly, namely to ·59 (p. e. ·15).

[1] Miss Smith, of Cherwell Hall, Oxford, who has since applied this test to women students and school girls, has also found a higher speed.

[2] The figures are the same for the 30 boys as for the 26 who alone sat for a 7th and 8th attempt, if decimals are disregarded. The corresponding figures for the Preparatory School are 63·8 (first tracing), 32·9 (sixth tracing).

[3] James, *Principles of Psychology*, Vol. I. p. 110.

Another defect in the form of the test adopted for these experiments revealed itself on plotting curves to represent graphically the course of the improvement in the various individuals; the character of these curves seems to indicate that the measurements on which they were based in some cases really measured a development due to at least two independent factors. Introspection suggests that besides the process of building up a new system of acquired co-ordinations between eye-movements and hand-movements by trial and error under the guidance of attention, there may also come into play simple kinaesthetic memory, i.e. the process of forming a disposition, purely motor, to move the hand over the pattern to be traced independently of visual guidance and of attention to the objects reflected in the mirror. In some subjects (not perhaps in children as young as those concerned in the present investigation) the movements may be actually directed by reflective inference as to the nature of the properties of reflection by plane mirrors, based either on previous knowledge or on observation during the course of the actual experiment.

Many of these defects might be remedied by further improvements in the form of the test, such for instance as the adoption of more irregular patterns to trace, and of different patterns of equal difficulty for the practice-attempts intervening between the first and last of a given sitting, and of a more adequate mathematical method of treating the experimental figures so yielded. It seems at any rate clear that in this test, apparently new, we have one both capable and worthy of further improvement and use[1].

[1] With a fresh set of 30 boys from the same school and also 12½ to 13½ years old at the time of experimentation, I have recently tried the following method of using the Mirror. The patterns contain 24 holes arranged in a circle 15 cm. in diameter. The holes are connected by straight lines forming chords of the circle, drawn two from each hole at irregular angles. Lines parallel to each other or to the edges of the card are avoided. The lines thus form an irregular zigzag route within the circle, returning on itself, and touching each hole upon the circumference once and once only. In pricking out the pattern the boy goes from hole to hole, not in order round the circle, but backwards and forwards across, according to the route thus traced by the lines. Each boy starts by pricking for the same period of time, namely two minutes; and the operator notes the number of holes pricked in this time. The subject then practises upon another similar pattern for five minutes. He then returns to the first pattern, pricks the same number of holes as on the first occasion, and his performance is now timed. The amount of improvement is measured by subtracting the time of this final performance from the time of the first, viz. from 120″; his retentiveness may be similarly measured. (Every boy to prick the same number of dots as the rest for his first attempt, and, for his final attempt, to prick for the same length of time as this given number originally took him, the additional number of dots then pricked to measure his improvement, would

(11) Spot Pattern.

Apparatus and Procedure.

The 'Spot Pattern' test was devised for the present purpose as a test of scope of apprehension by Mr McDougall; and is now being used by Dr Schuster as a mental test in the Anthropometric Laboratory at Oxford.

The chief piece of apparatus employed in this test is a Portable Tachistoscope, also devised by Mr McDougall. This consists of a vertical stand of wood attached to a horizontal base by a hinge,—which allows it to fold forward upon the base when not in use,—and by a detachable spring,—which keeps it upright when hooked to it at the back. In the upright stand is cut a circular aperture, about 7 cm. in diameter; and at the back of this is screwed a self-setting photographic time-shutter, made by Bausch and Lomb[1]; while the front—for ordinary tachistoscopic experiments—is covered with a semi-transparent card, sliding in grooves fixed on the face of the stand, and bearing printed on its hinder surface the object to be shown by means of the tachistoscope. The card is illuminated for a small fraction of a second by transmitted light from a lamp placed behind the stand, the duration of the exposure being regulated by the shutter.

In the Spot Pattern test the cards used were of a different nature. The objects to be shown consisted of patterns made by piercing 7, 8 or 9 holes in opaque squares of millboard; each card was covered with a sheet of white paper, pasted on the surface presented to the subject, so that the patterns could only be seen when light from behind was thrown upon them through the shutter. The nature of his task was first explained to the subject by reading to him a written form of instructions accompanied by a brief demonstration; and the test pattern was then shown to him in a dark room by means of the tachistoscope. The length of exposure chosen was $\frac{1}{25}$th sec. After seeing it five times at intervals of about $1\frac{1}{2}$ secs. the subject attempts to reproduce a copy of the pattern by mapping it on prepared sectional paper (see figure 1).

also embody the same principle, though the differentiation thus obtained is not quite so minute.) In this way, many of the defects mentioned in the text seem largely evaded; and the correlations with Intelligence prove to be slightly higher,—viz. improvability, ·73, retentiveness, ·48.

[1] A cheaper shutter is used by Dr Schuster, viz. a No. 6 Packard-Ideal Shutter with a $2\frac{1}{2} \times 4\frac{3}{4}$ in. aperture, set for instantaneous exposures.

The sectional paper is divided into large squares about the size of that figured below, each containing 36 (6 × 6) small squares; and the pattern to be copied is so pricked in the card that when reproduced the whole design will fall completely within the area of a single large square, while each of the spots of which it is composed will fall at the corner of a small square, i.e. at a point of intersection of two sectional lines. The advantage of this arrangement is that each of the spots in the boy's reproduction is definitely either right or wrong, and his performance is rendered susceptible of numerical evaluation. If the subject fails to reproduce the pattern correctly at the first attempt, he is again shown it five times as before, and then makes a second map in another large

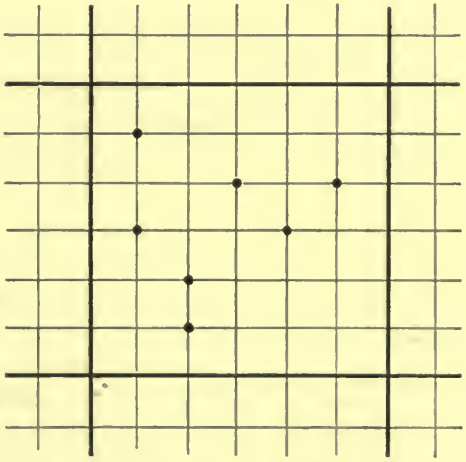

FIG. 1. Map of Spot Pattern as reproduced by subject.

square. This is repeated till the pattern is correctly reproduced. The total number of flashes or exposures required before he succeeds is taken as an inverse measure of the subjects' ability in this test, as shown by his proficiency with a pattern of stated difficulty, the difficulty of the pattern increasing with the number of spots, and with the irregularity of their arrangement[1].

Results.

The tachistoscope was found to require a larger amount of experience on the part of both subjects and operators than any of the other tests,

[1] For a brief account of the psychological nature of the process involved in such an experiment, cf. W. McDougall, *Physiological Psychology*, pp. 129, 130.

except perhaps those involving sensory discrimination. Six series were actually carried out at the Elementary School, four with patterns of seven spots, one with a pattern of eight, and one with a pattern of nine. The first series of all had to be rejected as worthless, owing partly to the irrelevant excitement aroused in the subjects by the 'electric flash' as the boys named it, and partly to the intervention of the holidays. The experiments with eight or nine spots were too difficult for many of the Elementary School boys, and as their inclusion slightly decreases the reliability coefficients and correlations they too have been omitted. There remain therefore three series, conducted successively by the writer, by Mr Flügel, and by the writer again; the reliability coefficient given by these is ·55. Three series were carried out at the Preparatory School, Mr Flügel using a card of seven spots, and the writer cards of seven and eight spots. At this school we were not able to obtain the complete darkness and silence procured at the other in our extemporised dark-room, and consequently the reliability coefficient and the raw correlations with Intelligence are not so high. The reliability coefficient here was ·50.

The element of unreliability in the separate series with this test seems very largely eliminated in the average of about three series; after this further experiments leave the averages much as before. The correlations of such amalgamated series with Intelligence, viz. ·76 and ·75 (p. e.'s ·05 and ·09) are fairly high and nearly the same at both schools. The average raw correlation at the Elementary School is, however, considerably higher than the average raw correlation at the Preparatory School; they measure ·64 and ·44 respectively; correcting to 1·00 and ·66 respectively.

The average number of exposures required to copy a pattern correctly by the boys of the Elementary group was 45·3 (M. V. 21·1; extremes 10 and 157; three sections, 19·5, 41, 71·6 respectively; weak-minded boy, 155 on one occasion, 220 on a second). In the Preparatory group the average was 38·6, where the series was made with an 8-spot pattern, since the 7-spot cards were found scarcely difficult enough to differentiate the better boys of this group (the M. V. being 12·5; the extremes 11 and 86·6); calculated on the basis of the series with 7-spot cards only, the group average for these boys is still better, namely, 30.

(12) Dotting Apparatus.

Apparatus and Procedure.

The apparatus here termed the 'Dotting Apparatus' is a machine for testing and graphically recording continued maximal voluntary concentration of attention. The method was devised by Mr McDougall; and an improved form of his apparatus has been suggested by Dr Rivers, and constructed by the Cambridge Scientific Instrument Company[1]. In an experiment conducted by means of this machine the task of the reagent is to mark with a pencil or stylographic pen[2] an irregular zigzag row of dots, lithographed in red upon a paper tape, carried past the field of view at an adjustable speed by a small wooden drum rotated by clockwork. The clockwork is driven by a heavy weight (some 9 kilos) pulling on a cord wound round a metal cylinder, which transmits, by means of a system of cog-wheels, a slow motion to the wooden drum and a faster motion to a specially designed friction-governor. The metal cylinder, cog-wheels and drum are covered by a low desk, which together with the weight is supported by a small table. In the lid of the desk is an oblong aperture or window, 5 × 10 cm., through which the dots are visible as the paper tape is carried past it by the wooden drum immediately beneath. The friction-governor is fixed at the side of the desk, accessible to the operator. The friction of a pair of revolving blunt brass points, pressed by the centrifugal force of a pair of weights against the rim of a metal disc around which they revolve, retards the motion of the machine. The amount of this friction is increased by increase in the rate of the motion of the machine (since this increases the centrifugal tendency of the weights), and thus automatically keeps the speed of the machine constant; the amount of friction is also regulated partly by the mass of the pair of weights, which are changeable at will and which thus provide a coarse adjustment of the speed; and

[1] The earlier form of the same apparatus is described in the *British Journal of Psychology*, Vol. II. No. 4, p. 435, W. McDougall, "On a New Method for the Study of Concurrent Mental Operations and of Mental Fatigue (Preliminary Communication)." A description of Dr Rivers' improvements will be found in his book on *The Influence of Alcohol and other Drugs on Fatigue*, Appendix II, p. 125.

[2] I venture to suggest that work with this apparatus would be yet more accurate if, instead of a pencil or stylo, a blunt mounted needle were used, and if the surface of the wheel or drum carrying the paper tape, instead of being hard and smooth, were covered with cloth or felt, to allow the needle to prick the paper. The considerations which led me to adopt the needle for the Tapping and Mirror tests (cf. pp. 132, 146) seem on consideration to apply with even greater force to the Dotting test.

partly by the pressure of a spring counteracting the pressure of the weights and regulated by a screw with a graded head, which thus provides a fine adjustment. By means of the fine adjustment the speed of the machine can gradually be changed by the operator, without stopping either the motion of the machine or the work of the reagent, from a speed of about 70 dots per minute to a speed of about 109 dots per minute with medium weights; with heavier or lighter weights similar ranges of speed can be obtained below or above these limits. The row of dots upon the paper band (see figure 2) is carefully designed so that the succession shall be as irregular as possible, the horizontal distance, however, of each dot from the last (i.e. the interval in the direction of motion) being always 5 mm., the extreme lateral deviation of the dots being 15 mm., and no dot deviating by more than 7 mm. from the line of its predecessor.

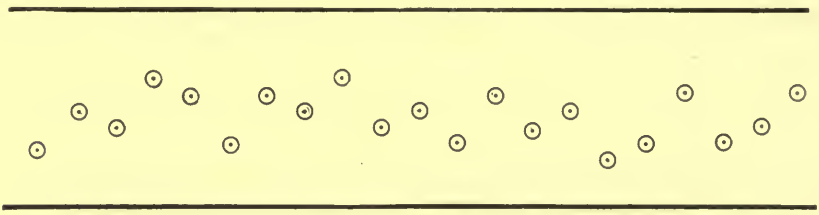

Fig. 2. Portion of Tape to be marked in Dotting test.

Sitting at the table, resting his wrist upon the desk, the reagent watches and marks the dots as they appear through the window, and are carried past towards his left. Each act of dotting constitutes a discrimination reaction, and a spell of dotting constitutes a series of such reactions performed at full, or nearly full, speed. Throughout the marking of a series moving at a given rate the subject's task is one of uniform difficulty, and the difficulty of the task depends upon the velocity at which the tape of paper moves. As he has the same fraction of a second for the accomplishment of each hit, he has to make a rhythmical series of strokes with the pencil; but as the position of each dot is unknown till it is seen, each stroke has to be aimed. This requires a sustained effort of attention, the degree of effort depending upon the rate of the rhythm of the strokes, and therefore measured by the rate of movement of the dots upon the paper tape. When marked, the paper furnishes a permanent graphic record of the maintenance of the effort, failure of continuity of attention being indicated by the presence of pencil-marks unaimed, or of red dots unmarked.

In the present investigation the measurement sought was the maximum rate at which each boy could mark the dots correctly. He was first given half a minute's practice, and then, after a short rest, commenced marking at the rate of about 80 dots per minute; this was continued for 10 secs., and, if the boy was fairly successful, the rate was increased to about 86 dots per minute; after another 10 secs. it was again increased by a similar increment, and so on till the boy completely failed to aim his strokes. His record was subsequently examined, and the speed at which he first made three or more omissions in 10 secs. (or their equivalent according to the system of marking unaimed hits, badly aimed hits, double hits, &c., suggested by Mr McDougall[1]) was accepted as the upper limit of his power of sustained attention. Where a boy went on for 50 secs., i.e. till he reached a speed of 110 dots per minute, he was given a second attempt beginning at a faster rate, as a higher range of speed is only attainable by stopping the machine, and altering the revolving weights of the friction-governor. This also evaded the highly fatiguing effects of a prolonged sitting.

Results.

It was not possible to obtain the Preparatory boys at the Laboratory (where this particular test was carried out) on more than one occasion, but that occasion was the same for every boy, viz. after morning school. Hence no reliability coefficient was obtainable for the Preparatory School, though the reliability was probably high. Each of the Elementary children was able to come twice. The reliability coefficient for the two series thus obtained for them was ·86. Part of the manipulation of the apparatus while the subject is actually at work devolves, it will be remembered, upon the superintendent of the experiment; so that the personal equation is in this test likely to be appreciable, though perhaps small; accordingly, as the series were all superintended by the same operator, the reliability coefficient is probably slightly too high. The Elementary group, partly owing to its size, was not able to work through the test at the same period of the day or week. The differences in fatigue this entailed were naturally slightly unfavourable to accurate work with an apparatus originally devised to test mental fatigue. Further, the right arms of three of the boys were either paralysed, recently broken, or recently dislocated. In consequence, the correlations with Intelligence are somewhat lower at the Elementary

[1] *Brit. Journ. of Psych. l.c.* p. 439.

School than at the Preparatory,—being ·60 at the one, and ·84 at the other, when calculated between the amalgamated series and the Headmaster's estimate; ·65 and ·84, when the raw correlations with the various estimates of Intelligence are averaged; and ·75 and ·96, when corrected by means of the Elementary School coefficient. In contrast with most of the other tests, little seems to be gained in this test by a repetition of the experiments.

The average speed[1] of the Elementary group was 12 (M. V. $1\frac{1}{2}$; extremes $17\frac{1}{2}$ and $10\frac{1}{4}$; average of the three sections $14\frac{1}{2}$, 13, $11\frac{3}{4}$; speed of the imbecile $8\frac{1}{2}$). The average speed of the Preparatory group was faster than that of the brightest section of Elementary boys, viz. 15 (M. V. 2; extremes 21 and 10).

As these figures suggest, the method of computing achievements at this test does not permit such thorough differentiation between every subject as do some of the other tests. It yields the largest portion of ties. This has necessarily lowered its coefficient of correlation with intelligence. Nevertheless, although the classification which this test can furnish is not minutely discriminative, yet its general character follows the order furnished by the masters with remarkably few contradictions. Since the average raw correlations for the separate series are as good as the correlations when these series are amalgamated into one, it is evident that the results of a single application of it may be relied upon without corroborative repetition.

Conclusions (Association and Attention Tests).

According to the corrected coefficients, these last four tests—Memory, Mirror, Spots and Dotting—have pure correlations with intelligence, on the whole, the highest of all. The Sensori-Motor tests, it is true, are close competitors with the Mirror and Memory; but the reliability coefficients of the Preparatory series, where the corrected correlations are so high as to raise the averages of the Sensori-Motor tests above those of the Mirror and Memory, are both below ·50. In general purpose, as well as in results, these last four tests seem to fall together. The Dotting test was devised specifically to test Attention. The three other tests have been classed as tests of Association. Each of them for

[1] The speed is given in terms of units of the machine, i.e. in terms of the scale upon the fine adjustment wheel with medium weights on the friction-governor (where the light or heavy weights were actually used, the necessary conversion of the observed readings has been made to keep the figures comparable). A speed of 12 represents a speed of about 90 dots a minute; 15 represents a speed of about 110 dots a minute.

most individuals must involve, as an integral part of the process whereby they are carried out, the building up, under the guidance of attention, of new associations between one set of stimuli and another, between one set of movements and another, or between stimuli and movements. Hence, in these tests also, Attention is a particularly essential factor. The fact that the results of the Sensori-Motor tests approximate so closely to the results of these last four Tests is readily explicable if the general account already given of the former be taken in conjunction with the general account here given of the latter.

GENERAL CONCLUSIONS.

THE DIAGNOSIS OF INTELLIGENCE.

We may now bring these results and inferences to bear upon the three main problems, with which, as noted at the commencement of the article, the experimental investigation of General Intelligence has to deal.

For practical deductions, the coefficients of the amalgamated series (and perhaps the average crude correlations) are of more significance than their theoretical corrections. Of the twelve tests, six furnish coefficients below ·50 and six furnish coefficients above ·50. The former six—the simple sensory and motor tests—are thus of little use in the empirical diagnosis of intelligence. Among the latter six, no single test, at any rate in its present form, can be claimed as a self-sufficient instrument for measuring and detecting ability in individuals. But they indicate the direction in which such a test may hopefully be sought. Particularly promising are the four new tests. Of these Mr McDougall's 'Dotting Machine' seems to be the most scientific. Where the external conditions of the experiment could be kept most uniform, namely with the small Preparatory group, both the amalgamated and the average raw coefficients reached ·84. Such uniformity is difficult in more extensive work, and the ensuing variety in attention and fatigue affect the performances with this test. Moreover, its figures are less discriminative than either of the other three. By increasing the number of spots in the patterns, the Tachistoscope test may be made to differentiate with almost any degree of minuteness. The apparatus, too, is portable. It is a slow test, however, and without repetition scarcely reliable. And it calls for some experience both on the part of the boys to grasp the nature of the task, and on the part of the experimenter to manipulate

the apparatus with regularity. These difficulties might perhaps be obviated by improving the instructions to be given to the boy, and the design of the cards shown, so as to make the technique simpler. Unlike the Dotting Machine and the Tachistoscope tests, the apparatus for the Mirror test can be procured with but little trouble or expense, and needs no trained superintendent. It too, however, requires further improvements, especially in procedure and calculations, to eliminate the influence of possible previous practice, and to elicit more completely the significance of the figures observed. If called upon to recommend a simple test for immediate use on untrained subjects, I should be inclined to advocate the Alphabet test as perhaps the simplest and most satisfactory of all.

Hitherto, we have considered each test in isolation. Let us now observe the effect of amalgamating the results of the best. On making a grand average of the six gradings from those that give coefficients above ·50 and arranging the boys accordingly, we obtain a list correlating with the Headmaster's order to the extent of ·85 at the Elementary School and ·91 at the Preparatory.

The lower correlation at the Elementary School is mainly due to five large discrepancies, of 5 to 8 places, between the amalgamated test order and the Headmaster's order. One of these is due to error in the test results, and is the case of a boy slightly incapacitated for them by a semi-paralysed right arm. One probably is due to overestimation by the master of a boy who, though in other tests inefficient, appeared to possess a remarkably good memory. The other discrepancies may possibly be due to a contrast effect which seems to have operated in the master's attempt to dovetail boys of different standards together; boys at the top of a lower standard have, I fancy, been placed too far above boys at the bottom of a higher standard. Most of these discrepancies in the Headmaster's list are corrected in the boy's and assistant master's lists (themselves not so reliable as the Headmaster's) in the same direction as the results of the tests. At one of the schools, the comparison of the final test results with those of the Headmaster drew his attention to a reticent boy who had apparently been allowed to take a lower position in the school than, according to the results of the tests, he appeared to deserve. This interesting point I owe to the kindness of the Headmaster himself, who subsequently told me that the boy was in consequence tentatively promoted and that this promotion was afterwards remarkably justified.

By means, then, of some half-dozen tests, we are able independently

to arrange a group of boys in an order of intelligence, which shall be decidedly more accurate than the order given by scholastic examinations, and probably more accurate than the order given by the master, based on personal intercourse during two or three years, and formulated with unusual labour, conscientiousness and care.

THE ANALYSIS OF INTELLIGENCE.

The results of the investigation, then, are so far positive; the empirical measurement of individual intelligence appears to be at least feasible. Towards the explanatory analysis of the general nature of intelligence their contribution may seem more indefinite.

Two questions are involved: Is any single explanation possible? if so, what is it? The only previous answers based on methods similar to the present are those of Dr Spearman's. His main conclusions were as follows: "Whenever branches of intellectual activity are at all dissimilar, then their correlations with one another appear wholly due to their being all variously saturated with some common fundamental Function (or group of Functions)"..."All examination in the different sensory, school, or other specific faculties may be considered as so many independently obtained estimates of the one great common Intellective Function." Hence "there exists a something we may provisionally term 'General Sensory Discrimination,' and similarly a 'General Intelligence,' and...the functional correspondence between these two is not appreciably less than absolute[1]."

The first of Dr Spearman's propositions, the "Theorem of the Universal Unity of Intellective Function" is tested by a corollary logically issuing from it, called that of the "Hierarchy of the Specific Intelligences[2]." Its principle may be most briefly expressed as follows:

$$\frac{r(A, P)}{r(B, P)} = \frac{r(A, Q)}{r(B, Q)},$$

where A, B, P, Q represent any four capacities not obviously akin[3].

Suppose, for instance, performance A correlates twice as highly as performance B with performance P; by hypothesis, this is because some

[1] *Loc. cit. supra*, pp. 72, 73. *Amer. J. Psychol.*, see however Dr Spearman's comment quoted below, p. 165, note 1.

[2] *Amer. J. Psychol. loc. cit.*

[3] This equation is immediately deducible from equation (*f*), *Zeitschr. f. Psychol. loc. cit. supra*, Vol. XLIV. p. 85.

central performance X, say discrimination, plays twice as great a *rôle* in A as in B. Then, for precisely the same reason, the correlation of A with a further performance Q will be twice as great as that of B with Q. Take now a series of functions, let us say reactions of varying complication, such as Dotting, Alphabet-sorting, Card-sorting, Card-dealing, Tapping, etc.; let us for illustration assume that these are specific manifestations of one common process, say Motor Co-ordination, more or less essential to all of them, and therefore connected with them in various degrees; then, if Dotting correlate with Card-sorting, Dealing, etc. in progressively diminishing degrees in that order, any other function of the same series, such as Alphabet-sorting, will also correlate with Card-sorting, Dealing, etc. in progressively diminishing degree in the same order; and similarly, if the correlation of Card-sorting with Dotting be higher than that of Card-sorting with Alphabet-sorting, then the correlation of Dealing with Dotting will also be higher than that of Dealing with Alphabet-sorting; so of Tapping, and similarly throughout the series in either direction. The system of correlations between each possible pair in such a series is called 'a hierarchy.'

The correlations between the various modes of investigating Intelligence described in these pages have been arranged as a hierarchy in Tables V and VI. The correlations are given in terms of the coefficients for the amalgamated series obtained at the two schools, since these are least open to criticism; analogous results hold good for average raw, corrected, and average corrected coefficients; hierarchies for these may readily be compiled from the complete list of figures at the end (p. 176). It will be found, however, that they are not so perfect. Dr Spearman and Prof. Krueger imply that satisfactory hierarchies are exhibited only by the 'pure' or theoretical coefficients; but it appears that those based on amalgamated measurements are better than those based on theoretical 'correction,' if the experimental conditions are carefully controlled. Each of the 13 tests, etc., furnishes 12 coefficients, one with each of the others; the sum of these 12 coefficients is taken as measuring its general tendency to correlate, and therefore provisionally determines its order in the hierarchy. The numbers are so printed that each shows the correlation between the function named vertically above it and that named horizontally to the left. Hence, if the functions measured are all specific forms of one common, fundamental function, then the values should always become successively smaller as the eye travels across to the right or downwards to the bottom.

TABLE V. *Hierarchy of Coefficients.* (*Amalgamated Series.*)

(A) Elementary School.

		Dotting apparatus	Alphabet	Sorting	Imputed intelligence	Dealing	Spot pattern	Tapping	Mirror	Sound	Lines	Touch	Memory	Weight
DOTTING APPARATUS	Observed coefficient	—	77	67	60	69	57	57	50	52	48	38	20	16
	Theoretical value ...	—	80	73	72	72	67	63	49	45	33	28	27	05
	Deviation	—	03	06	12	03	10	06	01	07	15	10	07	11
	P. e. of coefficient ...	—	05	07	08	06	08	08	09	09	09	11	12	12
ALPHABET	Observed coefficient	77	—	74	61	66	59	54	29	52	16	62	31	07
	Theoretical value ...	80	—	69	69	69	65	60	46	43	32	26	25	05
	Deviation	03	—	05	08	03	06	06	17	09	16	36	06	02
	P. e. of coefficient ...	05	—	06	08	07	08	09	11	09	12	07	10	12
SORTING	Observed coefficient	67	74	—	52	72	45	61	34	52	14	22	19	23
	Theoretical value ...	73	69	—	62	61	59	54	42	39	28	24	23	04
	Deviation	06	05	—	10	11	14	13	08	13	14	02	04	19
	P. e. of coefficient ...	07	06	—	09	06	10	08	11	09	12	11	10	19
IMPUTED INTELLIGENCE	Observed coefficient	60	61	52	—	44	76	47	67	40	29	13	57	-13
	Theoretical value ...	72	69	62	—	69	58	53	41	39	28	23	23	04
	Deviation	12	08	10	—	16	18	06	26	01	01	10	34	17
	P. e. of coefficient ...	08	08	09	—	10	05	10	07	10	08	12	08	12
DEALING	Observed coefficient	69	66	72	44	—	76	47	67	40	29	13	57	-13
	Theoretical value ...	72	69	61	60	—	58	53	41	39	28	23	23	04
	Deviation	03	02	11	16	—	07	12	01	05	19	00	04	03
	P. e. of coefficient ...	06	07	06	09	—	10	07	11	12	10	11	12	12
SPOT PATTERN	Observed coefficient	57	59	45	76	51	—	41	41	47	25	03	26	11
	Theoretical value ...	67	65	59	58	58	—	48	37	35	35	26	21	04
	Deviation	10	06	14	16	07	—	07	04	12	01	18	05	07
	P. e. of coefficient ...	08	08	09	05	09	—	10	10	10	11	12	11	12
TAPPING	Observed coefficient	57	53	61	47	65	41	—	41	47	08	26	-05	22
	Theoretical value ...	63	60	54	53	53	48	—	36	34	25	20	20	04
	Deviation	06	06	07	08	12	07	—	05	13	18	06	25	18
	P. e. of coefficient ...	08	09	08	10	08	10	—	10	10	12	11	12	12
MIRROR	Observed coefficient	50	29	34	67	40	45	45	—	34	16	08	05	-05
	Theoretical value ...	49	46	42	41	41	37	36	—	25	19	15	15	03
	Deviation	01	17	08	26	01	04	05	—	09	03	07	10	08
	P. e. of coefficient ...	09	11	11	17	10	10	10	—	10	12	12	12	12
SOUND	Observed coefficient	52	52	52	40	34	47	47	34	—	-07	-01	01	-13
	Theoretical value ...	45	43	39	39	39	35	34	25	—	17	14	14	02
	Deviation	07	09	13	01	05	12	13	09	—	24	15	13	15
	P. e. of coefficient ...	09	09	09	10	17	10	10	12	—	12	13	12	12
LINES	Observed coefficient	48	16	14	29	47	25	08	16	-07	—	26	06	19
	Theoretical value ...	33	32	28	28	28	26	26	25	17	—	10	10	02
	Deviation	15	16	14	01	19	01	17	03	24	—	16	04	17
	P. e. of coefficient ...	09	12	12	08	10	11	12	12	12	—	11	12	12
TOUCH	Observed coefficient	38	62	22	13	23	03	26	08	-01	26	—	16	29
	Theoretical value ...	28	26	24	23	23	21	20	15	14	10	—	08	01
	Deviation	10	36	02	10	00	18	06	07	15	16	—	08	28
	P. e. of coefficient ...	11	07	12	12	12	12	11	12	12	11	—	12	11
MEMORY	Observed coefficient	20	31	19	57	19	26	-05	05	01	06	16	—	05
	Theoretical value ...	27	25	23	23	23	21	20	15	12	10	18	—	01
	Deviation	07	06	04	34	04	05	25	10	13	04	08	—	04
	P. e. of coefficient ...	12	10	11	10	12	11	12	12	12	12	12	—	12
WEIGHT	Observed coefficient	16	07	23	-13	01	11	22	-05	-13	19	29	05	—
	Theoretical value ...	05	05	04	04	04	04	04	04	03	03	01	01	—
	Deviation	11	02	19	17	03	07	18	08	15	17	28	04	—
	P. e. of coefficient ...	12	12	12	12	12	12	11	12	12	12	11	12	—

Average deviation = ·100. Average p. e. = ·101.

TABLE VI. *Hierarchy of Coefficients.* (*Amalgamated Series.*)

(B) Preparatory School.

		Dotting apparatus	Alphabet	Imputed intelligence	Mirror	Memory	Spot pattern	Tapping	Sorting	Sound	Lines	Weight	Touch	Dealing
Dotting Apparatus	Observed coefficient	—	84	84	71	69	62	48	73	48	25	07	03	-03
	Theoretical value ...	—	85	80	76	70	66	66	60	48	39	14	-07	-13
	Deviation	—	01	04	05	01	04	16	13	00	14	07	10	10
	P. e. of coefficient ...	—	06	06	10	12	12	16	10	16	19	20	20	20
Alphabet	Observed coefficient	84	—	80	48	84	67	57	76	34	22	-14	-28	45
	Theoretical value ...	85	—	78	74	68	64	64	58	51	37	14	-07	-12
	Deviation	01	—	02	26	16	03	07	18	17	15	28	21	57
	P. e. of coefficient ...	06	—	07	16	06	15	14	09	18	19	20	19	16
Imputed Intelligence	Observed coefficient	84	80	—	54	78	75	43	56	37	17	-19	-06	29
	Theoretical value ...	80	78	—	70	64	60	60	55	44	35	13	-06	-12
	Deviation	04	02	—	16	14	15	17	01	07	18	32	00	41
	P. e. of coefficient ...	06	07	—	14	08	09	16	14	17	20	19	20	18
Mirror	Observed coefficient	71	48	54	—	43	38	75	34	57	54	44	31	-44
	Theoretical value ...	76	74	70	—	61	58	57	52	42	34	12	-06	-11
	Deviation	05	26	16	—	18	20	18	18	15	20	32	37	33
	P. e. of coefficient ...	10	16	14	—	16	17	09	18	14	14	16	18	16
Memory	Observed coefficient	69	84	78	43	—	74	54	64	17	28	-05	-35	03
	Theoretical value ...	70	68	64	61	—	53	53	48	39	31	11	-06	-10
	Deviation	01	16	14	18	—	21	01	16	22	03	16	29	13
	P. e. of coefficient ...	11	16	18	16	—	09	14	11	20	19	20	18	20
Spot Pattern	Observed coefficient	62	67	75	38	74	—	38	51	25	34	07	-44	19
	Theoretical value ...	66	64	60	58	53	—	50	45	36	29	11	-05	-10
	Deviation	04	03	15	20	21	—	12	06	11	05	04	39	29
	P. e. of coefficient ..	12	15	09	17	09	—	17	15	19	18	20	16	19
Tapping	Observed coefficient	48	57	43	75	54	38	—	48	28	44	34	07	-31
	Theoretical value ...	66	64	60	57	53	50	—	45	36	29	11	-05	-09
	Deviation	16	07	17	18	01	12	—	03	08	15	23	12	22
	P. e. of coefficient ...	16	14	16	09	14	17	—	16	19	17	18	20	19
Sorting	Observed coefficient	73	76	56	34	64	51	48	—	38	00	-22	-14	02
	Theoretical value ...	60	58	55	52	48	45	45	—	33	27	10	-05	-09
	Deviation	13	18	01	18	16	06	03	—	05	27	32	09	11
	P. e. of coefficient ...	09	08	14	18	11	15	16	—	17	20	19	16	20
Sound	Observed coefficient	48	34	37	57	17	25	28	38	—	07	34	17	-17
	Theoretical value ...	48	51	44	42	39	36	36	33	—	21	08	-04	-07
	Deviation	00	17	07	15	22	11	08	05	—	14	26	21	10
	P. e. of coefficient ...	16	18	17	14	20	19	19	17	—	20	19	20	20
Lines	Observed coefficient	25	22	17	54	28	34	44	00	07	—	35	19	-13
	Theoretical value ...	39	37	35	34	31	29	29	27	21	—	06	-03	-06
	Deviation	14	15	18	20	03	05	15	27	14	—	29	22	07
	P. e. of coefficient ...	19	19	20	14	19	18	17	20	20	—	18	19	20
Weight	Observed coefficient	07	-14	-10	44	-05	07	34	-22	34	35	—	38	-35
	Theoretical value ...	14	14	13	12	11	11	11	10	08	06	—	-01	-02
	Deviation	07	28	32	32	16	04	23	32	26	29	—	39	33
	P. e. of coefficient ...	20	20	19	16	20	20	18	19	19	18	—	17	18
Touch	Observed coefficient	03	-28	-06	31	-35	-44	07	-14	17	19	38	—	-48
	Theoretical value ...	-07	-07	-06	-06	-06	-05	-05	-05	-04	-03	-01	—	01
	Deviation	10	21	00	37	29	39	12	09	21	22	39	—	49
	P. e. of coefficient ...	20	19	20	18	18	16	20	16	20	19	17	—	15
Dealing	Observed coefficient	-03	45	29	-44	03	19	-31	02	-17	-13	-35	-48	—
	Theoretical value ...	-13	-12	-12	-11	-10	-10	-09	-09	-07	-06	02	01	—
	Deviation	10	57	41	33	13	29	22	11	10	07	33	49	—
	P. e. of coefficient ...	20	16	18	16	20	19	19	20	20	20	18	15	—

Average deviation = ·165. Average p. e. = ·162.

How far, then, do these observed correlations form an ideal hierarchy? They clearly do not fit into the proposed scheme with perfect precision. Nor indeed can we expect them to. Like all empirical observations they are subject to error. What we have to demand is the following.

Firstly, their deviations from the ideal hierarchy should, on the whole, be neither more nor less than the probable erroneousness of the observations. The theoretical values for the ideal hierarchy may be obtained by various mathematical formulae[1]. These are given in the tables, together with the deviations of the observed coefficients from them, and the probable errors for the coefficients given in two decimal places only. For the Elementary School group the average deviation works out at ·100, while the average probable error comes to ·101. For the Preparatory School group the average deviation works out at ·165, while the average probable error comes to ·162. So far, then, a neater agreement between observation and theory could scarcely be desired[2].

Hardly less reassuring is the accordance disclosed on turning from the average deviation to the extreme deviations. In a 'normal' chance distribution, we should expect a deviation three times greater than the probable error to occur about once in 24 times. Here we have 78 coefficients for each group. Here, then, we should expect such a deviation to occur about three times in each. Actually it occurs four times in the Elementary School, and twice in the other. Again, a deviation

[1] The following simple formula has been supplied for this purpose by Dr Spearman (to whom I am here particularly indebted for several improvements on my own demonstration of a hierarchy):

Let $r(s, t)$ denote the required theoretical value, satisfying the condition

$$\frac{r(A, P)}{r(B, P)} = \frac{r(A, Q)}{r(B, Q)},$$

and at the same time according as well as possible with the correlations actually observed. Then

$$r(s, t) = m_s \cdot m_t,$$

where

$$m_s = \frac{a_s}{\sqrt{2\Sigma - a_s}} \cdot \sqrt{\frac{n-2}{n-1-\frac{n \cdot a_s}{2\Sigma}}},$$

a_s = the sum of all the correlations with the performance s,

Σ = the sum of all the different correlations altogether,

n = the number of performances,

and m_t has a value analogous to m_s.

[2] For perfect mathematical strictness in comparing the probable errors and the average deviations, certain corrections should first be made. These, however, are complicated, and leave the results virtually the same as before; they have consequently here been omitted.

four times as great as the probable error may be expected to occur by mere chance about once in 124 times. Here it occurs twice in the Elementary School, and not at all in the other. Some of these deviations are themselves suggestive. At the Elementary School, three of the four large deviations occur with Imputed Intelligence, namely in its correlations with the Spot Pattern, Mirror, and Memory tests; such irregularities are here quite natural, since the method of estimating intelligence was not homologous with the methods of estimating the other capacities. The other correlations of intelligence may in consequence very likely have been reduced. The largest augmentation occurs with Memory; and towards Memory the schoolmaster's estimate of Intelligence is, as we have already seen, specially liable to be biassed. The deviation of the Alphabet-Touch correlation appears to be merely 'accidental.' At the Preparatory School, the deviations of Dealing are probably due to the exaggeration in its negative coefficients. The deviation of the Alphabet in its correlation with Memory might be explained by the fact that these two tests both involved rapid recognition of printed letters.

The tendency for subordinate groups of allied tests to correlate together is discernible, but small. At the Preparatory School there are distinct signs that the sensory tests are specially connected; the correlations of Weight with Sound, with Lines and with Touch are above ·30 instead of under ·10, and those of Touch with Sound and with Lines are over ·15 instead of being negative. At the Elementary School the correlations of Weight with Lines and with Touch, and of Touch with Lines, are again somewhat high; but the deviations are never more than twice the probable error. Among motor tests, at the Elementary School Dealing correlates with Tapping and with Sorting a little more highly than is demanded by the hierarchy, but at the Preparatory School the two corresponding coefficients are respectively negative and approximately zero. As to the sensori-motor tests, Sorting at both schools correlates with Alphabet a little highly for its place, though only at the Preparatory School is the difference greater than the probable error. All these special deviations are under three times the probable error. In this respect, the results of a wider application of the Test-methods confirm and extend the observations of Dr Spearman, who found in his own experiments "the range of the central function... so universal, and that of the specific functions so vanishingly minute."

The main significance of this hierarchy of experimental performances, is, as it appears to me, that we are led to infer that all the functions of

the human mind, the simplest and most complicated alike, are probably processes within a single system. A process typical of higher psychophysical 'levels' may be connected with a process typical of lower psychophysical 'levels' far less intimately than either is with a process of intermediate 'levels.' Yet this relatively small correlation is not a disproof, but a consequence of, their inclusive organisation within a single integrative system of psychical dispositions or of neural arcs. The contrary assumption of a radical dichotomy between "the general mammalian foundation of the central nervous system" and the "specifically human capacity" of General Intelligence,—towards which Dr Archdall Reid[1], and even Professor Thorndike[2] seem to incline,—proves a serious barrier to the advance of the biological standpoint in individual psychology.

The nature of the general or fundamental function cannot here be discussed with any hope of finality. The methods employed must first be extended to yet higher levels of mental process. The process of Abstraction was indirectly involved in one of the memory tests. But there results were negative. With this exception, the highest mental levels—the conceptual and relational—have been left by our experiments absolutely untouched.

It is clear, however, from our extension of the tests into regions representing a stage of mental development higher than sense-perception, that the absolute identification of General Intelligence and General Sensory Discrimination (if it has ever been suggested by any but its opponents[3]) cannot be maintained. The sensory tests fall together in the hierarchy, and the principle common to them, which determines both their falling together, and the order into which they fall, appears to be in the main but a specific mode or determination of a wider principle which determines the order of the whole series: for, as we have just seen, the generic or hierarchical tendency has far more

[1] See quotations *inf.* p. 170.

[2] See his discussion of Dr Spearman's hypothesis, *Amer. J. Psychol.*, July 1909, "Relation of Accuracy in Sensory Discrimination to General Intelligence," from which the phrases cited in the text are borrowed.

[3] With reference to my criticism of the passage cited above (p. 159) formulating his view of the relation of General Sensory Discrimination and General Intelligence, Dr Spearman has written to me: "This conclusion of mine was badly worded. I did not mean (as others have naturally taken it) that general intelligence was based on sensory discrimination; if anything, *vice versa*. I take both the sensory discrimination and the manifestations leading a teacher to impute general intelligence to be based on some deeper fundamental cause, as sketched in the *Zeitschrift für Psychologie*, Vol. XLI. p. 110, para. 5."

influence upon the size of the sensory coefficients than any specific or extra-hierarchical tendencies for sensory tests to correlate particularly highly together. Thus the correlations between pairs of sensory processes, like those between all other pairs, "appear" mainly, if not "wholly due to their being all variously saturated with some common fundamental Function," which also saturates the rest of the processes tested, and determines the correlations between all other pairs. This, however, does not prove that the highest common factor of all processes correlated with Intelligence is identical with the *highest* common factor in sensory discriminations: but only that the former is a necessary factor, though not necessarily the sole, perhaps even a vanishingly minute factor in the latter.

One important universal condition of all experiments upon school-children is Goodwill, that is, Interest as determining Zeal. Some such factor seems often to have contributed to influence differentially the psychological measurements obtained from untrained subjects in previous investigations, and it might plausibly be supposed that Goodwill was the main condition permeating the hierarchy of coefficients set forth above, and determining its order. If so, the order of tests in the hierarchy should correspond to the degree in which Interest or Zeal was displayed in the subjects. Lists have been obtained from representative boys, arranging the tests, as far as possible, in the order in which they 'liked' or 'disliked' them. This order in the average also tallies with the degree of Zeal which the tests seemed to the experimenter to evoke in the children. It is,—beginning with the test in which interest was keenest: Mirror, Card-dealing, Card-sorting, Spot Pattern, Sound, Tapping, Alphabet, Weight, Dotting machine, Touch, Lines, Memory. There is little difference between the orders given by the boys of both schools. None of the lists throw light on the hierarchies. We have, therefore, to look for some other common factor besides Goodwill in the sense of Zeal determined by Interest.

The test which correlates most with all the other tests, and consequently heads the hierarchy, is the Dotting test. The Dotting test was specially devised to measure power of sustained effort of maximal concentration, in short to test Voluntary Attention. The inference is that the power of Voluntary Attention is the capacity, common to all the functions tested, which enters most into the processes involved. The hypothesis that Attention is the essential factor in Intelligence is already a well-known one. In view of it, before the hierarchies were drawn up, the tests were arranged in order, according to the degree in

which Attention might be expected to be required in the successful performance of the tasks. Such arrangements were obtained from interrogations of the boys, and independently from two or three psychologists. The average arrangement is as follows: Dotting, Spot Pattern, Memory, Mirror, Alphabet, Sorting, Sound, Lines, Touch, Weight, Dealing, Tapping. This corresponds fairly closely with the order of the various correlations with Intelligence, and nearly as closely with the orders given by the hierarchies. The most interesting discrepancies are those of the Alphabet test and the Memory test. In both hierarchies the Alphabet test stands second to the Dotting test; it may be that the high correlations of Alphabet-finding with Imputed Intelligence and tests of Intelligence, point to the connection between quickness to recognise letters and readiness to think in terms of written words, which in civilised societies forms an important element in rational intelligence. The fact that Imputed Intelligence does not head the list, but follows Dotting and Alphabet, need not, on the present hypothesis, cause surprise; for, if Intelligence essentially involve some central factor such as Attention, the schoolmaster's opinion may be an impure and indirect method of estimating what is estimated far more directly and scientifically by these experimental tests. The displacement of Memory is very suggestive. In the hierarchy for the Preparatory School, it takes its place along with the 'association' tests. In the hierarchy for the Elementary School, it falls in among the sensory tests. An observation of the manner of the boys when performing this test, and an inspection of their papers, suggest that the Elementary boys relied for the most part upon a primary memory image of the words, while the Preparatory boys relied rather upon associations formed between them. Thus, instead of reproducing the words in the correct order, the Elementary boys often wrote down the last two or three as they were ringing in their ears, or reproduced the column in inverted order, and, in endeavouring to complete the number of words, commonly waited for the missing word to "recur spontaneously"; whereas the Preparatory boys seldom wrote the words in reversed or inaccurate sequence, and, on forgetting, muttered the words they had retained in their proper order, in the hope of the chain of associations suggesting the missing link. Thus, so far as memory implies mere retentiveness of sensory images, it seems to bear little relation to intelligence; so far as memory implies organisation of new associations, it seems to bear a high relation to intelligence.

This point, taken in conjunction with several stray inferences noticed

under the conclusions as to the several groups of tests[1], strongly suggests that it is one feature or function of attentive consciousness in particular, which forms the basis of Intelligence—namely, the power of re-adjustment to relatively novel situations by organising new psycho-physical co-ordinations. A comparison of the successive correlations with Intelligence yielded by the same test, where it was repeated with the same reagents on different occasions, seems to lend support to this suggestion. The differences are most marked in the simple correlation of each of the two or three series with the Headmaster's provisional estimate. These are given in Table VII. In the average correlations with *all* the provisional estimates the differences are still present, and nearly always similarly but less strikingly shown. (It is the average of all these correlations, not merely of the coefficients given below, that furnishes the 'average raw correlation.')

TABLE VII. *Intelligence coefficients for successive series with the same test.*

	Elementary School						Preparatory School		
No. of series	1st	2nd	3rd	4th	5th	6th	1st	2nd	3rd
Dotting apparatus	·67	·62					·84		
Spot pattern[2]	·19	·57	·58	·69	·64	·61	·27	·64	·47
Mirror	·51	·22					·54		
Memory	·57	·38	·46				·81	·71	
Alphabet	·53	·45	·38				·83	·65	·45
Sorting	·44	·29	·56				·43	·71	
Dealing	·51	·42	·46				·13	– ·03	
Tapping	·23	·44					·43		
Sound	·50	·43					·37		
Lines	·28	·17					– ·04	·54	
Touch	·25	– ·05	·17				·03	– ·25	
Weight	·06	– ·14	·03				·17	– ·08	

It will be noticed that (with one or two exceptions, of which the Spot Pattern is the most prominent) the correlation is highest on the first occasion, that is to say, when the task is newest. This is not due to a difference of operator, since the first set of experiments was not in every case undertaken by the same operator; and further where the operator taking the first set took also the third, the correlation for his series is (with the above exceptions) always higher than that for his later series. The suggestion is that this reduction of correlation is due to reduction in Attention. Whether this in turn is due to decrease of

[1] See pp. 136, 137, 144. [2] See p. 152.

interest and zeal, or to mechanization of the subject's general procedure; and whether the subsequent partial recovery is due to the greater interval elapsing before the third series, or to the personal equation of the operator, it would be premature here to speculate.

We have seen throughout that the greater the change, and the greater the complexity, and the greater the novelty involved in the task performed, the greater also (*ceteris paribus*) is the Imputed Intelligence of the performer. To relative novelty all the other attributes are probably secondary. Thus high intelligence seems to mean high capacity for continually systematising mental behaviour by forming new psycho-physical co-ordinations, older co-ordinations being retained, so that newer co-ordinations bring with them increased complexity and incessant change. In such progressively integrative actions of the mind the efficient and directive agent is attentive consciousness. And in this sense we may agree that so-called 'Voluntary' Attention is, of all recognised psychological processes, the essential factor in General Intelligence. This interpretation may help to bring the general use of the term Intelligence into relation with the special significance it has acquired in animal psychology. It further suggests that we may eventually seek the psycho-physical basis, underlying this capacity, in a particular characteristic of general neural constitution; the accentuation of such a neural characteristic would then produce the type of mind known as intelligent, while its biological inheritance would form the condition of the transmissibility of the mental trait.

THE INHERITANCE OF INTELLIGENCE.

The third and last phase of our problem is in many ways the most important of all. The gathering interest in "the possible improvement of the human breed," the growing belief that the innate characters of the family are more potent in evolution than the acquired characters of the individual, the gradual apprehension that unsupplemented humanitarianism and philanthropy may be suspending the natural elimination of the unfit stocks—these features of contemporary sociology make the question whether ability is inherited one of fundamental moment.

Hitherto, writers have dogmatically assumed and asserted, some—the existence of class differences in native intelligence, others—the absence of such differences, all without any sure basis for their opinion.

Supporting their existence is the authority of the first and foremost exponent of Eugenics, Sir Francis Galton. His investigations of '*Hereditary Genius*' have, however, dealt rather with types of ability peculiar in nature, or exceptional in degree. The best known statistical enquiry into the transmission of ordinary mental faculties is that of Professor Karl Pearson. The essence of his method was roughly as follows: a large number of schoolmasters and schoolmistresses were consulted about pairs of brothers and sisters in their schools, and were asked to fill up a statement of their opinion of the ability, temper, popularity, etc., of these various pairs. The evidence thus accumulated on collateral transmission was strongly affirmative, and Prof. Pearson states as his main conclusion "that the mental characters in man are inherited in precisely the same way as the physical." Now, to discover that, in social, moral, and intellectual nature, children of the same family resemble one another, suggests that they owe it to a common parent, but does not prove that its reappearance is due to biological inheritance, since mental characteristics—especially as they appear to the casual observer—may with equal probability have been handed down by training and tradition, and re-acquired by each succeeding generation through imitation and habit. Prof. Pearson's results are, therefore, inconclusive.

As an authority for the opposite view I may quote Dr Archdall Reid. His argument is, "supposing a child of refined and educated English parents were reared from birth by African cannibals. Then, in body, when grown, the child would resemble his progenitors more than his captors, but does anyone believe the same of his mind?...The common sense of mankind has universally recognised this radical difference between mind and body," namely, that the capacities of the mind can be, and should be, trained, the capacities of the body may be left to develop of themselves. From such general considerations he infers that "the evidence is overwhelming that the mental and moral qualities are not inherited in the same sense as the physical qualities." He adds as the explanation that, in addition to instincts like those of the lower animals, man "has an enormous memory, and enormous power of utilising its contents." ("...The two combined make up what we term intelligence....") "It is this educability that confers on man all his morality, all his intelligence, all his intellectuality, all his reasoning power, all his adaptability....The instincts of men have everywhere been the same, for instincts are inherited in the same sense as physical characters are inherited. But man's knowledge, aspirations, ideals, and

all that flows from them, belong to a different and to a higher category. They are acquirements, and as such are not inherited by offspring." And, simply "because no man's experiences are quite the same as those of any other man, individual men of the same family or class differ widely amongst themselves; men of different classes differ yet more, and men of different nations even yet more."

This explanation, however, really concedes all that Eugenics could desire. For this "educability" is obviously a mental character; it is also admittedly inherited. It is consequently quite possible that in different families such educability may itself be inherited in different amounts. To determine whether such inherited differences actually exist, and, if so, what is their relation and their proportion to acquired mental differences,—these are questions which special experiment alone can finally decide. For, if there are important innate differences, then (in Dr Archdall Reid's own words) "these are so completely masked and overshadowed by immensely more important acquired differences that they cannot be recognised without much closer scientific investigation than has yet been attempted[1]."

There is here, then, a legitimate, definite and urgent problem. Inductive generalisation and deductive inference are alike inconclusive. Experimental evidence is imperative. Meagre though they be, therefore, the indications afforded by the foregoing results may justifiably be exploited.

The investigation was concerned with two representative groups. Each of these was fairly homogeneous; each represented a distinct social type. The boys of the Elementary School were drawn from the lower middle classes. Their parents were chiefly small tradesmen and artisans, and, in most cases, presumably neither pre-eminent nor defective in intelligence. The parents of the Preparatory boys were, as we have seen, mostly persons occupying high positions in the ecclesiastical, civil, or academic world—positions implying eminence in ability and culture. The class represented by them is a product of a highly differential social selection, perhaps the most efficient form of selection now operating in our society. Whether the difference of intelligence in the parents was itself inherited, or acquired, we have no need to ask. We have only to observe whether the difference reappears in the children, and whether the reappearance is to be attributed to direct inheritance. The experimental data furnished by the application of the 12 tests to

[1] *Sociological Papers*, Vol. III. pp. 93, 94.

TABLE VIII.

	Elementary School								Preparatory School			
	Av.	M.V.	Best	Worst	Clever	Ordinary	Dull	Imbecile	Av.	M.V.	Best	Worst
Dotting Apparatus (1 unit = about 7 dots per min.)	12	1½	17	10¼	14½	13	11¼	8½	15	2	21	10
Spot Pattern (number of exposures per pattern)	45·3	21·1	10	157	19·5	41	71·6	157	38·6	12·5	11	86·6
Mirror (number of secs. per 6 tracings)	388·9	90·1	181	747	347	383·7	429·8	2464	257	54	156	526
Memory (4 marks = 1 correct word)	134·2	27·7	223·3	68·6	167	135	134·2	—	216	29	265·5	153·5
Alphabet-finding (number of secs. per 26 letters)	91	15·4	50	138	75·6	90·5	103	306	74	13	48·5	119
Card-sorting (number of secs. per 50 cards)	48·1	4·1	34	61·6	41	49·3	51·5	87	41·3	4·1	36·5	49·5
Card-dealing (number of secs. per 50 cards)	31·6	5·9	20·3	46	26	31·6	36·6	46	27·3	1·8	20	34
Tapping (number of taps per 15 secs.)	80·5	8	97	59	86·5	80·2	76·3	—	95·5	11·9	119	66
Sound (threshold in vibration differences)	6	1·9	1	11·5	4·6	5·6	7·1	10	3·5	2·2	·3	8
Lines (A. C. E. in cms.)	·41	·107	·20	1·05	·36	·36	·50	—	·39	·179	·07	1·30
Touch (threshold in mms.)	36·2	9	19	58·3	31·3	37·3	38·3	20	38·9	11	12·5	63·7
Weight (threshold in gms.)	8·75	1·5	6	16	8·5	9·5	8·1	4	9·3	1·6	5	11·5

the two respective groups have been brought together in Table VIII. The first column of figures gives the average measurement for each test at the Elementary School group. The individuals' Mean Variation about the average follows. Next are given the best and the worst individual performances in the same group, and then the averages for the three sub-groups—namely the clever, the ordinary, and the dull or infra-normal boys respectively. For convenience of reference the achievements of the weak-minded boy are also added. The last four columns give the group-average, with its Mean Variation, and the optimum and pessimum, for each test at the Preparatory.

With two exceptions the average performances of the boys of the Preparatory School are all superior to those of the boys of the Elementary School. The two tests in which they are equal to, or weaker than, the latter are those for Discrimination of Weight and Discrimination of Two Points upon the Skin, and these two are the only two tests which yielded negative correlations with Intelligence. Hence, wherever there are correlations with Intelligence, there (so far as we can discover) boys of superior parentage are themselves superior. Moreover, at the Sound, Tapping, Memory, Mirror, Alphabet, and Dotting tests, the Preparatory boys are superior even to the cleverest section of the Elementary boys.

Now in the case of the lowest social classes, general inferiority at mental tests might be attributable to unfortunate environmental and post-natal influences, such as improper nourishment, improper air, and, generally, to neglected conditions of physical health. But (with the possible exception of the dullest boy of the 30) such conditions could not be suspected with the boys who, at a fee of 9*d.* a week, attended the Central Elementary School. Again, unlike superiority in scholastic examinations, superiority in the experimental tests could not have depended, wholly or in part, upon any previous acquisition of material. Nor yet could it in every case have depended upon previous special practice. In the Mirror and Memory tests no doubt special practice played an important part in raising the proficiency of the Preparatory boys. In Pitch-discrimination, Musical training, and in Tapping (perhaps also in Alphabet-finding and Dotting) motor training might be suggested as the determining factors; but we have seen that the opportunities for musical training were probably greater with the Elementary boys; and the special facilities for manual training at the Elementary School probably developed motor rapidity and motor co-ordination as much as the special facilities for games of muscular skill

at the Preparatory School; moreover, the superiority is not uniformly conspicuous in motor tests as such: it is nearly as marked in Sorting, but much less marked in Dealing. Nor can the Preparatory boys' achievements be attributed entirely to superior training generally. School education is immediately concerned with certain special departments of learning; and operates primarily by developing certain special interests and aptitudes. So far as such education may also be conceived as indirectly exercising the mind generally for *all* activities into which intelligence may subsequently enter, the training in general intelligence (in the sense in which the term has here been used) would probably be considered by most authorities to be as good as the best Elementary School of an academic city, as at a private school preparing its scholars for examinations of a specific kind. As implied above, the former type of school from its very nature tends rather to train its scholars to use their minds, while the latter tends rather to train its scholars to store their minds. Hence, so far as scholastic education is concerned, the boys of the Elementary School seem on the whole to have been better equipped to perform the tests than the others, if there was any appreciable difference of equipment in this respect at all. There remains the educative influence of home and social life,—of the wider experience and more intellectual atmosphere enjoyed by the Preparatory boys as compared with the Elementary boys. Important for mental development as this influence undoubtedly is, I yet believe (as my remarks on the personal attitude of the reagents, and upon the effect of zeal and interest must have suggested) the part played by this factor in the tests was small. Here, however, one must confess, such speculative arguments can convey little conviction to those who have not witnessed the actual manner of the respective boys; perhaps the following experimental data may, therefore, carry more weight.

Firstly, we seem to have, for one important instance, a direct estimate of the influence of such irrelevant factors. Among a group of some 60 boys from the Elementary School, 10 to 16 years old, proficiency in the Alphabet test was found to correlate with Age to an extent of ·29. Now Age, at any rate during this particular period, in a group otherwise fairly homogeneous, may be taken as a simple and direct means of estimating the extent to which post-natal, or environmental, factors, likely to influence intelligence generally, have operated in the several boys. The older the boy, the greater the amount of training his mind must have acquired from individual experience. At the same time, his intelligence may also have undergone a greater

spontaneous development. So that the coefficient of correlation between Alphabet and Age indicates that the connexion between Alphabet and training through general experience is no more, and probably less, than ·29. Hence, compared with the connexion between Alphabet and Intelligence, this connexion is negligible.

A second piece of evidence is afforded by a series of experiments carried out upon the Elementary boys 18 months after the several series already described. In September, 1909, Mr Flügel and myself revisited the school, and repeated the more important tests upon as many of our original subjects as were still at the school, or could still be traced[1]. The results were unexpected. It had been anticipated that the new measurements might show suggestive relations to the new occupations or the new schools of the boys who had left the Elementary school, or to the progress in class of the boys who remained. On the contrary, the new series show nearly as high correlations with earlier series, as the earlier series did with each other; and the measurements for the boys tested remained in most cases much as before. The data may be summarised most briefly by expressing the improvement or deterioration for the set of boys re-tested as a percentage of the average calculated for the same set of boys from the original experiments. Touch shows a deterioration of 3 °/₀; Comparing Lines an improvement of 8 °/₀; Dealing an improvement of 15 °/₀; Card-sorting an improvement of 6 °/₀; Alphabet-sorting an improvement of 4 °/₀; Memory a deterioration of 9 °/₀; the Spot Pattern a deterioration of 7 °/₀; Dotting a deterioration of 3 °/₀; the Mirror alone shows a marked improvement, viz. 31 °/₀, and this is doubtless attributable to the persistence of the practice gained during the earlier experiments. With this exception, the capacities re-tested seem during the interval to have remained all but stationary. Yet the boys' mental equipment has not. A somewhat 'dull' boy, for instance, who was 25th in the amalgamated list for Six Tests in 1908, has in 1909 risen to a place in the school occupied in 1908 by a 'clever' boy, who was then 4th on the amalgamated list; yet his new measurements, instead of concomitantly rising to equal those of the 'clever' boy, are now equivalent to those of the boy who was 24th. Similarly with most of the other boys. Thus, though the period between the ages of 13 and 15 is for boys one of rapid progress in knowledge,

[1] I am much indebted to the employers of some of the boys for their readiness to allow them to be tested during hours of employment, and to furnish interesting information as to the intelligence of the boys concerned.

interests, and acquired aptitudes, yet in the capacities measured by the tests no corresponding alteration is made. Hence, these capacities appear to constitute a relatively permanent endowment; and consequently it seems legitimate to assume that they depend upon innate differences in the individuals concerned.

For all these various reasons we may conclude that the superior proficiency at Intelligence tests on the part of the boys of superior parentage, was inborn. And thus we seem to have proved marked inheritability in the case of a mental character of the highest "civic worth."

Parental intelligence, therefore, may be inherited, individual intelligence measured, and general intelligence analysed; and they can be analysed, measured, and inherited to a degree which few psychologists have hitherto legitimately ventured to maintain.

TABLE IX. *Complete List of Correlations.*

	Reliability coefficient		Coefficient of amalgamated series		Av. raw coefficient		Corrected coefficient		
	Elem. Sch.	Prep. Sch.	Elem. Sch.	Prep. Sch.	Elem. Sch.	Prep. Sch.	Elem. Sch.	Prep. Sch.	Av.
Imputed Intelligence...	·88	·91	—	—	—	—	—	—	—
and Dotting apparatus...	·86	—	·60	·84	·65	·84	·75	·96	·85
Spot pattern	·55	·50	·76	·75	·61	·44	1·00	·66	·83
Mirror	·52	—	·67	·54	·50	·47	·74	·68	·71
Memory (Total)	·70	·93	·57	·78	·48	·76	·60	·82	·71
,, Concrete ...	—	—	·58	·84	—	—	—	—	—
,, Abstract ...	—	—	·48	·78	—	—	—	—	—
,, Nonsense...	—	—	·43	·75	—	—	—	—	—
Alphabet	·60	·48	·61	·80	·50	·61	·68	·91	·79
Sorting	·84	·38	·52	·56	·45	·60	·53	1·00	·76
Dealing	·88	·80	·44	·29	·48	·05	·54	·06	·30
Tapping	·51	—	·47	·43	·44	·28	·65	·41	·53
Sound	·67	—	·40	·37	·40	·35	·52	·41	·46
Lines	·50	·54	·29	·17	·34	·31	·51	·44	·47
Touch	·73	·75	·13	– ·06	·14	– ·14	·17	– ·17	·00
Weight	·86	·51	– ·13	– ·19	– ·01	– ·14	– ·01	– ·20	– ·10
Dotting Apparatus and Spot pattern ...			·57	·62	·52	·40	·80	·62	·71
Mirror.....			·50	·71	·55	·71	·84	1·00	·92
Memory			·20	·69	·17	·73	·22	·84	·53
Alphabet			·77	·84	·61	·54	·85	·84	·84
Sorting			·67	·73	·64	·57	·75	1·00	·87
Dealing			·69	– ·03	·72	– ·14	·83	– ·17	·33
Tapping			·57	·48	·55	·48	·83	·73	·78
Sound.....			·52	·48	·52	·48	·68	·63	·65
Lines			·48	·25	·40	·30	·60	·45	·52
Touch.....			·38	·03	·22	·25	·30	·31	·30
Weight			·16	·07	·06	·16	·07	·25	·16

	Coefficient of amalgamated series		Av. raw coefficient		Corrected coefficient		
	Elem. Sch.	Prep. Sch.	Elem. Sch.	Prep. Sch.	Elem. Sch.	Prep. Sch.	Av.
SPOT PATTERN and Mirror	·41	·38	·40	·37	·75	·75	·75
Memory	·26	·74	·25	·55	·41	·84	·62
Alphabet	·59	·67	·56	·41	·96	·83	·89
Sorting	·45	·51	·47	·20	·68	·47	·57
Dealing	·51	·19	·43	·25	·66	·40	·53
Tapping	·41	·38	·34	·25	·64	·50	·57
Sound	·47	·25	·34	·23	·55	·40	·47
Lines	·25	·34	·16	·44	·50	·85	·67
Touch	·03	– ·44	·13	– ·17	·21	– ·29	– ·04
Weight	·11	·07	·19	·05	·27	·10	·18
MIRROR and Memory	·05	·43	·08	·44	·13	·64	·38
Alphabet	·29	·48	·39	·28	·71	·56	·63
Sorting	·34	·34	·37	·32	·64	·71	·67
Dealing	·40	– ·44	·48	– ·40	·72	– ·61	·05
Tapping	·41	·75	·25	·74	·48	1·00	·74
Sound	·34	·57	·24	·55	·40	·93	·66
Lines	·16	·54	·08	·50	·16	·94	·55
Touch	·08	·31	·17	·25	·27	·40	·38
Weight	– ·05	·44	– ·06	·37	– ·10	·71	·30
MEMORY and Alphabet	·31	·84	·28	·69	·47	1·00	·73
Sorting	·19	·64	·19	·52	·27	·90	·58
Dealing	·19	·03	·14	·05	·18	·06	12
Tapping	– ·05	·54	·01	·52	·01	·80	·40
Sound	·01	·17	·13	·20	·19	·27	·23
Lines	·06	·28	·03	·19	·05	·27	·16
Touch	·16	– ·35	·11	– ·23	·15	– ·27	– ·06
Weight	·05	– ·05	·05	·15	·07	·22	·14
ALPHABET and Sorting	·74	·76	·59	·60	·83	1·00	·91
Dealing	·66	·45	·61	·23	·83	·37	·60
Tapping	·54	·57	·37	·47	·67	·94	·80
Sound	·52	·34	·43	·20	·68	·35	·51
Lines	·16	·22	·20	·17	·37	·32	·35
Touch	·62	– ·28	·44	– ·29	·66	– ·48	·09
Weight	·07	– ·14	·19	– ·03	·26	– ·06	·10
SORTING and Dealing	·72	·02	·66	·19	·77	·32	·54
Tapping	·61	·48	·52	·43	·78	1·00	·89
Sound	·52	·38	·19	– ·32	·25	– ·62	– ·18
Lines	·14	·00	·21	·03	·32	·07	·19
Touch	·22	– ·14	·17	– ·17	·20	– ·32	– ·06
Weight	·23	– ·22	·20	– ·19	·23	– ·33	– ·05
DEALING and Tapping	·65	– ·31	·52	·23	·79	·36	·57
Sound	·32	– ·17	·32	– ·06	·41	– ·06	·17
Lines	·47	– ·13	·38	– ·05	·59	– ·07	·26
Touch	·23	– ·48	·26	– ·37	·32	– ·51	– ·08
Weight	·01	– ·35	·05	– ·26	·05	– ·40	– ·17
TAPPING and Sound	·47	·28	·31	·26	·53	·44	·48
Lines	·08	·44	·08	·30	·16	·57	·36
Touch	·26	·07	·07	·06	·12	·09	·10
Weight	·22	·34	·13	·38	·11	·74	·42
SOUND and Lines	– ·07	·07	– ·03	·14	– ·05	·23	·09
Touch	– ·01	·17	– ·01	·19	– ·01	·26	·12
Weight	– ·13	·34	·05	·31	·06	·53	·29
LINES and Touch	·26	·19	·16	·13	·26	·21	·23
Weight	·19	·35	– ·11	·08	– ·16	·15	·00
TOUCH and Weight	·29	·38	·29	·38	·37	·61	·49