THE EVIDENCE FOR THE CONCEPT OF INTELLIGENCE

BY CYRIL BURT

I.—THE NON-STATISTICAL EVIDENCE: (1) observational; (2) biological; (3) physiological; (4) individual psychology.

II.—THE STATISTICAL EVIDENCE: (1) the general factor; (2) the factor as cognitive; (3) the factor as innate—the hypothesis of multi-factorial inheritance.

III.—SUMMARY.

IV.—REFERENCES.

I.—THE NON-STATISTICAL EVIDENCE.

Current Criticisms.—The concept of intelligence, and the attempt to measure intelligence by standardized tests, have of late furnished a target for vigorous attack. The objections urged are partly practical and partly theoretical. Yet few of the critics show a clear or correct understanding of what the term really designates or of the reasons that have led to its introduction. Two misconceptions have become widely current.

(i) Those writers who are chiefly interested in the more practical issues, like Dr. Heim and Dr. Blackburn, explain that intelligence “is a popular and relatively unambiguous word,” and denotes a quality that “all can recognize, though few can define.” It follows that, instead of pinning I.Q.s. on to the coat of each child, we should leave any decisions that may be necessary to the intuitive insight of the teacher. Unfortunately, in a vain effort to measure the immeasurable, the modern psychologist “has been induced to restrict the meaning of the term to a vague quantitative abstraction.” No two of them, however, agree as to how that abstraction is to be defined. Hence “those who go chasing this ignis fatuus get quickly bogged down in mathematical abstruseness.” Meanwhile, the layman, so Mr. Richmond assures us, has begun to “sense a certain absurdity in measuring something called ‘intelligence’ without knowing what that something is or how it is defined.”

(ii) Those who are concerned with the more technical aspects of the subject apparently suppose that the concept was invented by a small band of statistical enthusiasts—Dr. Kirman (13) mentions Spearman, Pearson, and myself—who deduced their theories by primitive factorial procedures that have since been “publicly discredited.” The more accurate methods of Thurstone and his American followers, it is said, have since clearly shown that the intellectual achievements of different individuals are the product, not of a single general factor, but of a number of more specialized ‘primary abilities.’ And this at once accounts for the difficulties that beset all attempts to define intelligence. As Captain Kettle observed, when asked why the pictures of the Saghalien sea-serpent showed such incredible differences: “I’ve seen it’s because there’s no such crittur’; so each just draws his own fancy.”

The Definition of Intelligence.—Now the critics who protest about “the spate of incongruous definitions” usually rest their complaint on the results of the famous Symposium organized some twenty years ago. The Editor of an


2 (18), p. 227. Similar criticisms have also been put forward by Dr. E. G. Chambers, Dr. D. H. Stott, and Dr. C. M. Fleming.

3 For a recent statement of the American view, see A. Anastasi: Psychological Testing (1955), pp. 15, 33f.

4 Symposium on Intelligence and its Measurement,” J. Educ. Psych., XII, 1921, pp. 123-147 and 195-216. In framing his question, the Editor specifically asked, not how is intelligence to be defined, but “what do you conceive intelligence to be, and how can it best be measured? Should the test material call into play analytical and higher thought processes, or should it deal rather with simple, with associative, or with perceptual processes, etc.?”
American journal submitted two searching questions about the nature of intelligence to a dozen different psychologists, and received a dozen different replies. But the varying descriptions suggested were not, as Dr. Heim and others have supposed, intended to be ‘definitions’ in the strict logical sense: they were, in the language of J. S. Mill, merely ‘attempts to explain the thing,’ not ‘attempts to interpret the word.’ As the editorial letter shows, the purpose of the discussion was primarily a practical one—to determine how intelligence appears to operate, with a view to ascertaining ‘what material may most profitably be used in constructing tests.’ But that is quite a separate question, and except incidentally will not concern us here. Nor shall I discuss the validity of mental measurement or the practical value of the I.Q.—problems that are continually confused with the fundamental issue. The questions I now want to settle are prior to all these, namely, (i) how precisely should the term be defined, and (ii) what evidence is there for believing that something really exists corresponding to the definition proposed? However, instead of taking the term for granted and hunting round for a plausible formula, as is most frequently done, a sound scientific procedure requires us to start with the relevant facts. Let us, therefore, take the second of our two questions first.

History of the Concept.—Many of the criticisms to which I have alluded spring largely from a manifest ignorance as to how the concept originated. A rapid glance at the literature is, therefore, needed first of all. As a brief historical review will show, long before the advent of statistical analysis, several converging lines of evidence had already drawn attention to an important property of the mind, for which some special name seemed desirable. How its nature was envisaged can best be gathered by recalling the actual statements of leading authorities in each field.

(1) Observational.

The earliest attempts to analyse and classify the activities of the mind were based partly on the observation of various types of person in everyday life and partly on introspection. Plato, to whom we owe the basic distinctions, draws a clear contrast between ‘nature’ and ‘nurture’ (φύσις and τροφή); and then distinguishes three parts or aspects of the soul—τὸ λογιστικὸν, ἁπλὴν καὶ θυμικόν (Republic, 435A.). The modern terms—intellectual, emotional, and moral, cognition, affection, and conation—suggest rough but somewhat inexact equivalents for these untranslatable expressions. In a celebrated passage (Phaedrus, 253D) he sketches a picturesque analogy which conveys a better notion of the fundamental difference: the first component he compares to a charioteer who holds the reins, and the other two to a pair of horses who draw the vehicle; the former guides, the latter supply the power; the former is the cybernetic element, the latter the dynamic.

Aristotle makes a further contribution of lasting importance. He

---

1 For a discussion of these questions I may refer to Professor Vernon’s address on ‘The Psychology of Intelligence and G’ in the current Bulletin of the Brit. Psychol. Society (No. 20, pp. 1-14), which I had not seen before this article was written.

2 A more detailed account will be found in my ‘Historical Sketch,’ which forms the first chapter of the Board of Education Report on Psychological Tests of Educable Capacity (2, pp. 1-61) and in a recent Galton Lecture on ‘The Meaning and Assessment of Intelligence’ (5). The antecedent evidence, drawn from the four main fields reviewed below, was briefly summarized in my earliest papers on general intelligence (e.g., J. Exp. Pedag., 1. 1911, pp. 96). If the reader refers to that article, he will see that the criticism made by Dr. Maberley, and repeated in varying terms by several later writers—namely, that I ‘claimed to deduce the general factor from a statistical analysis of test-data’—quite misrepresents my argument: the statistical analysis was intended merely to confirm a hypothesis reached on far more concrete grounds.
contrasts the actual or concrete activity with the hypothetical 'capacity' on which it depends (δύναμις), and thus introduces the idea of an 'ability.' Plato's threefold classification he reduces to a twofold. For him the main distinction is between what he calls the 'dianoetic' (cognitive or intellectual) capacities of the mind and the 'oricetic' (emotional and moral).

Finally, Cicero, in an endeavour to supply a Latin terminology for Greek philosophy, translates δύναμις by facultas, and ὄρετικ by appetitio or sometimes conatus; while to designate διανοω he coins a new word, rendering the Greek term almost literally by the compound 'intelligentia.'

Here then we have the origin of both the concept and the term. So far from being a 'word of popular speech,' whose meaning has been restricted and distorted by the modern psychologist, intelligence is a highly technical expression invented to denote a highly technical abstraction. From Aristotle and Cicero it descended to the mediaeval schoolmen; and the scholastic theories in turn became elaborated into the cut-and-dried schemes of the faculty psychologists and their phrenological followers.

(2) Biological.

As Guilford has reminded us, the modern notion of 'intelligence as a unitary entity' was 'a gift to psychology from biology through the instrumentality of Herbert Spencer.' Following Aristotle and the later Scottish school, Spencer recognizes two main aspects of mental life—the cognitive and the affective. All cognition (he explains) involves both an analytic or discriminative and a synthetic or integrative process; and its essential function is to enable the organism to adjust itself more effectively to a complex and ever-changing environment. During the evolution of the animal kingdom, and during the growth of the individual child, the fundamental capacity of cognition 'progressively differentiates into a hierarchy of more specialized abilities'—sensory, perceptual, associative, and relational, much as the trunk of a tree sprouts into boughs, branches, and twigs. To designate the basic characteristic he revives the term 'intelligence.'

Evidence favouring Spencer's somewhat speculative theories was adduced by Romanes, Lloyd Morgan, and other pioneers of comparative psychology; and his views on intelligence were accepted, not only by British biologists like Darwin, but also by continental writers, like Binet and Claparède. Certainly, Mendel's earliest disciples maintained that the doctrine of unit-characters was utterly irreconcilable with the inheritability of a graded trait, such as intelligence (cf. 6, pp. 333f.); but, as we shall see in a moment, the later developments of the Mendelian hypothesis not only permit it, but actually suggest it.

(3) Physiological.

The clinical work of Hughlings Jackson, the experimental investigations of Sherrington, and the microscopical studies of the brain carried out by Campbell, Brodmann, and others, have done much to confirm Spencer's theory of a

1 De Anima, II, 3, 414a, 31. Eth. Nic., 1, 13, 18, 1102b, 30. The usual rendering 'power' must not be taken to imply causal agency: Aristotle is simply describing what Professor Broad has called a 'dispositional property.'

2 H. Spencer: Principles of Psychology (1870). I have summarized Spencer's views more fully in a recent article ('The Differentiation of Intellectual Ability,' this Journal, XXIV, 1954, pp. 76f).

3 Cf. C. Darwin: The Descent of Man (1888), I, pp. 101f.; G. J. Romanes: Animal Intelligence (1890); and Lloyd Morgan: Animal Life and Intelligence (1798).
'hierarchy of neural functions,' with a basic type of activity developing by fairly definite stages into higher and more specialized forms. In particular, the examination of the cortex, both in mental defectives and in normal persons, suggests that the quality of the nervous tissue in any given individual tends to be predominantly the same throughout. Defectives, for example, exhibit a "general cerebral immaturity"; their nerve-cells tend to be "visibly deficient in number, branching, and regularity of arrangement in every part of the cortex." After all, as Sherrington himself points out, much the same is true of almost every tissue of which the human frame is composed—of a man's skin, bones, hair, or muscles: each is of the same general character all over the body, although minor local variations are usually discernible. In the adult human brain marked differences in the architecture of different areas and of different cell-layers are perceptible under the microscope; but these specializations appear and develop progressively during the early months of infant life. And, of course, such differentiation is precisely what the Spencerian theory would entail.

The experimental study of the brain leads to the same conclusion. The intact brain acts always as a whole. No part of the brain functions in total isolation from the rest, as the older champions of cortical localization originally assumed. The activity, in Sherrington's phrase, is "patterned not indifferently diffuse"; but the patterning itself "involves and implies integration." Lashley's conclusions about the 'mass action' of the brain seem to lend further corroboration to that view; and, as several writers have suggested, this 'mass-action' might well be identified with $g$.

The evidence of neurology, therefore, itself suggests something very like a theory of general ability, which gradually differentiates into more specific functions, though we must beware of picturing such functions as separate 'faculties' located in certain centres or compartments of the brain, after the fashion of the older phrenologists and of several recent writers on so-called 'physiological' or 'medical' psychology.

(4) Individual Psychology.

All these earlier writers were interested primarily in the working of the mind as such, that is to say, in problems of general psychology. The first to apply scientific methods to the problems of individual psychology was Galton. Darwin and Spencer had maintained that the basic capacities of the human mind were hereditary, transmitted as part of our common racial endowment. Galton went farther and maintained that individual differences in these capacities were also innate. As a result of his investigations into 'hereditary genius,' he was led to discard the traditional explanation in terms of faculties and types, and to substitute a classification in terms of 'general ability' and 'special aptitudes':

---


2 J. S. Bolton: The Brain in Health and Disease (1914).

3 K. S. Lashley: Brain Mechanisms and Intelligence (1929). The experiments of Lashley and his colleagues consisted in training animals to perform definite tasks, and then removing parts of their brains: the animals were then re-tested, and in some instances re-trained. The main conclusion was that ability to learn depends, not so much on the nature or location of the tissue remaining, but upon its amount.

4 This identification is suggested by Sherrington (Man on His Nature, 1940, p. 288). It should be added that the details of Lashley's conclusions are not entirely free from criticism; but here we are concerned only with the major principle.
Evidence for the Concept of Intelligence

of the two he considered that general ability was "by far the most powerful".1 The differences between individuals formed, so he believed, not a set of distinct and discontinuous classes, as the type-theory assumed, but a series of continuously varying gradations, distributed more or less in accordance with the normal curve, i.e., much like differences in head-length, arm-length, or stature (10, pp. 23f., 35f.).

The Definition Implied.—These converging lines of inquiry, therefore, furnished strong presumptive evidence for a mental trait of fundamental importance defined by three verifiable attributes: first, it is a general quality; it enters into every form of mental activity; secondly, it is (in a broad sense of the word) an intellectual quality—that is, it characterizes the cognitive rather than the affective or conative aspects of conscious behaviour; thirdly, it is inherited or at least innate; differences in its strength or amount are due to differences in the individual's genetic constitution. We thus arrive at the concept of an innate, general, cognitive ability. We cannot, however, keep repeating a cumbersome phrase of twelve syllables every time we wish to mention it. And, since a name that suggests its own meaning seems preferable to a brand-new esoteric symbol, what better label can be found than the traditional term 'intelligence'?2

Here then is a clearly formulated hypothesis, the outcome of centuries of shrewd observation and plausible conjecture—a psychological hypothesis fully in accord with the findings of the biologist and neurologist. Nevertheless, each of the three propositions that I have just laid down has been vigorously challenged; and each has started off a protracted controversy that still remains unresolved.

At this point, therefore, the need for ad hoc inquiries based on rigorous statistical analysis becomes obvious. It is the function of statistical procedures to decide between alternative hypotheses by testing their verifiable corollaries. The claim of the factorist is not, as his critics so often imagine, to 'discover' mental abilities, running round with a cry of 'Eureka' whenever he has extracted a fresh factor: his object is merely to confirm or refute certain hypothetical concepts or components that have been tentatively reached on more concrete grounds. Let us then take each of the three foregoing propositions in turn, and consider what evidence, if any, is provided by these more cogent techniques.

1 Many contemporary writers, particularly in the field of education, attribute the antithesis between 'general' and 'special' abilities to Spearman, and identify it with the contrast between what he called g and s. Spearman himself, however, frankly admits that his own theories were prompted by those of Galton and Spencer. However, in his earlier papers he eventually rejected the notion of 'special aptitudes,' as merely a relic of the discredited faculties of the older school: the only 'specific' capacities that he recognized were those 'specific' to each particular test (cf. Amer. J. Psych., XV, 1904, pp. 74f, and 206f, and 20, pp. 6f).
2 In educational psychology the popularity of the term is due to the work of Alfred Binet, himself an avowed follower of Spencer and Galton. Like Galton, Binet firmly believed in the existence of a 'general ability,' and repeatedly distinguished it from what he called 'partial aptitudes.' This ability, he says, enters into 'nearly all the phenomena with which the experimental psychologist has previously concerned himself—sensation, perception, memory, as well as reasoning,' i.e., it is essentially a cognitive capacity. Finally, he explains that his intelligence tests were deliberately constructed to measure innate differences, in contrast to his pedagogical tests which measure acquired attainments (cf. esp., L'Année Psychologique, XI, 1905, pp. 191f., 245f.).

Galton himself more frequently spoke of 'general ability.' But at times he used 'intelligence' as a synonym, especially when the context called for the adjective (e.g., 9, p. 336). Those who fear the ambiguities of the more familiar term can use a literal symbol: I have suggested using γ for the hypothetical quality defined as above, and keeping g for the empirical measurement, with a subscript to indicate the method of measurement.
The logic of the argument should be carefully noted. In the natural sciences a direct deductive proof is out of the question: the mode of proof must be indirect and inductive. Hence, the conclusions reached can never be certain, but only probable. The critic commonly misses this point. He revels in demonstrating that some alternative interpretation can readily be conceived. But one can always think up alternatives. The verdict must depend on determining and balancing the crucial facts. A probable hypothesis can only be overthrown by showing that its rival is still more probable. And equally, of course, the defender of a hypothesis must prove that every alternative that is worth considering is less probable than his own.

II.—THE STATISTICAL EVIDENCE.

(1) The General Factor.

At the beginning of the century, the problem which chiefly exercised students of individual psychology was, in Bain's phrase, "the classification of intellectual abilities or powers." (i) Were there, as the faculty psychologists maintained, a number of specialized abilities, each independent of the rest—observation, practical ability, memory, language, reasoning, and the like? (ii) Or was there, as Ward maintained, "not a congeries of faculties, but only a single subjective activity"—a general capacity for cognition as such? (iii) Were there, as Galton believed, both a general ability and a number of more or less specialized capacities? (iv) Or, finally, might there be, as the earlier associationists and most of the later behaviourists alleged, no discernible structure in the mind at all?

Each hypothesis entailed its own distinctive corollaries; and Galton's technique of correlation offered a ready-made method of checking them. Thus, the obvious plan for attacking such a many-sided issue was to devise and apply suitable tests for the main forms of mental activity, and then calculate the correlations between each test and the rest. If, for example, the orthodox behaviourist is right, and there is "no organized structure in the mind—no ground for classifying mental performances under one or more broad headings, no basis for inferring efficiency in one type of activity from efficiency in another," then we should expect all the intercorrelations to be zero or at least non-significant.1

1 Spearman, writing of the "momentous investigation by Cattell and Wissler"—the first to apply "the Galton-Pearson coefficient of correlation" to the results obtained with psychological tests—evidently understands them to have accepted this inference (20, p. 56). Wissler, it is true, says that at first sight the low coefficients suggest that "every act measured by the tests is special and unrelated to every other act" (22, p. 55): but he plainly does not intend this conclusion to be final: he speaks of a "deep conviction that we are otherwise constituted," and points out that certain correlations (e.g., for memory and College grades) are positive and significant. Thorndike also said it was tempting to infer from the data that "there is nothing whatever (his italics) common to all mental functions or to any part of them" (Amer. J. Psych., XX, 1909, p. 368): but he, too, quickly abandoned this view. The reasons for the low correlations obtained in these earlier researches are now quite clear: (a) the earlier tests had a low reliability; (b) the functions tested were extremely simple, and the size of the correlation tends to increase with the complexity of the function; (c) the groups tested (students or school classes rather than complete age groups) were already highly selected for general intelligence.

Thomson's sampling theory, though expressed in language similar to that of the 'anti-structural psychologists,' leads to very different corollaries. "The Mind," he says, "has little structure: unlike the body, it is not sub-divided into distinct organs, but forms a comparatively undifferentiated complex of innumerable elements." These he pictures as 'bonds,' i.e., interconnecting neural paths: they have the same character or quality throughout the brain. But, so far from the effects of specific stimuli being limited to specific neural paths (as the earlier opponents of structure assumed), "any sample whatever of these elements can be assembled in the activity called for by a 'test'" (21, pp. 303, 306). Now
If, on the other hand, the mind consists of a number of specialized faculties or abilities, such as 'observation' (assessed by tests of sensory capacity) or 'practical ability' (assessed by tests of motor capacity), then we should expect that all the inter-correlations between the sensory tests would be positive and similarly that all the inter-correlations between the motor tests would be positive; on the other hand, we should expect that all the cross-correlations between the one group and the other would be approximately zero. If what Thorndike called 'the theory of natural compensation' held good, then the cross-correlations would actually become negative, since the 'sensory type' would be deficient in the characteristic capacities of the 'motor type' and vice versa. Lastly, if there were no specific faculties at all, but only 'a single cognitive activity'—'attention,' as Ward believed, 'sensory discrimination' as Sully maintained—then we should expect the entire table of correlations to exhibit what Spearman called a 'perfect hierarchical order,' or (in the more precise language of the mathematical textbook) to have 'a rank of one'—apart, of course, from minor aberrations due to sampling errors.

The results of the earlier inquiries revealed, almost without exception, positive and significant correlations between every form of cognitive activity. This disproves hypotheses (i) and (iv). Further, except when the sample was small and the sampling errors large, there were nearly always well-marked clusters of augmented correlations confined to similar forms of cognitive activity, and leaving significant residuals after the general factor was removed. This rules out hypothesis (ii). We are thus left with hypothesis (iii) as the only alternative consistent with the facts. And, accordingly, the unavoidable inference is that both a 'general factor' and a number of 'group factors' must be at work.1

But we are not yet justified in identifying this abstract 'general factor' with anything so concrete as 'general intelligence.' In Spearman's investigations 'general intelligence' is always represented by an external criterion, i.e., either by direct assessments for intelligence as popularly understood or (in later researches) by standard tests, selected as furnishing accredited 'reference values.' In my own investigations, the 'general cognitive factor' forms an internal criterion, namely, what I called the 'highest common factor' in the battery of tests. And to determine the concrete nature of such a factor, or this (as Thomson recognizes) is merely another version of the general factor theory: the chief difference is that with Spearman the general factor is identified with something concrete (mental energy); with Thomson it represents something abstract (the fact that the neural elements have the same general character throughout). The corollaries are plain. First, since 'the physical body has an obvious structure,' the contribution of the general physical factor should be much smaller for correlations between bodily measurements than for correlations between mental; indeed, it was this supposed contrast with physical measurements that led Thomson to promulgate his theory. Secondly, with mental measurements, the correlation table, even if not as completely hierarchical as Spearman believed, ought always to exhibit a 'low rank.' Recent work has falsified both these corollaries. To begin with, in the very table for physical measurements which Thomson cites, the contribution of the general factor is practically the same as for mental measurements (50 per cent. or rather more, Brit. J. Psych., Stat. Sec., II, p. 116); secondly, the application of mental tests to much larger samples shows that the low rank of the tables Thomson has in mind resulted from the small numbers tested, whereas the physical measurements were obtained from 3,000 persons. It may be added that no neurologist would subscribe to the view that a stimulus, whether simple or complex, merely 'sampled' the neural elements: the responses to the simpler stimuli are relatively specific and selective; the response to more complex stimulation essentially involves the integration or organization of the neural elements.

rather of the processes that give rise to it, a supplementary investigation is requisite, based on observations or introspections, or on the correlation of the factor measurements with independent gradings.¹

Later investigators, notably Brown, Thomson, and more recently Thurstone, have argued that, if we accept the existence of group factors or 'primary abilities,' we can dispense with the hypothesis of a general factor by assuming that the group factors overlap. But this solution has proved unworkable both in theory and in practice. When the general factor accounts for much more of the variance than any single group factor, or indeed than all the group factors put together, there is no theoretical gain in closing one's eyes to its presence. And in educational practice the rash assumption that the general factor has at length been demolished has done much to sanction the impracticable idea that, in classifying children according to their varying capacities, we need no longer consider their degree of general ability, and have only to allot them to schools of different types according to their special aptitudes; in short, that the examination at eleven plus can best be run on the principle of the caucus-race in Wonderland, where everybody wins and each gets some kind of prize.²

In their more recent writings, most of the opponents of the 'general factor hypothesis' have, more or less openly, withdrawn their opposition. Brown, for example, ultimately acknowledged that "the evidence for a general factor now seems conclusive." Thomson himself has constructed numerous booklets for testing intelligence. And Thurstone has proposed a scheme of 'second order factors' which shall expressly include a 'general factor' and so account for the correlations between the 'first order factors' or 'primary abilities.'³

(2) The Factor as 'Cognitive.'

Merely to demonstrate the presence of a general factor common to all cognitive activities does not (as is usually assumed) prove that this factor is specifically cognitive. One might as well argue that, because a general factor can be demonstrated common to all sensory activities, therefore this factor is simply and solely a capacity for sensory discrimination. Impressed by this obvious fallacy, a number of writers went on to argue that in all probability the factor common to mental and scholastic activities was not cognitive but conative. Such an interpretation had a warm appeal for those who cherished the doctrine of intellectual equality. When a pupil lagged behindhand in nearly

¹ Actually teachers' gradings for 'intelligence' (as I showed in my 1909 research) are markedly biased in favour of memory or capacity to learn; and many psychologists (e.g., Colvin) adopted this as a definition of intelligence. Spearman, following Sully and the sensationalist school, originally equated intelligence with 'sensory discrimination,' as the basic form of mental analysis. Ward, Stout, and others inclined to identify it with 'attention' or 'apparception,' i.e., mental or 'neotic' synthesis. This early disagreement about the nature of intelligence is no reason for repudiating the concept: after all, there is little agreement about the nature of gravity; but that is no reason for discarding the principle. And, in point of fact, the conflict can easily be reconciled if we borrow the suggestion of the neurologists and suppose its function to be that of 'integration,' i.e., organization (which involves both analysis and synthesis).² For a fuller discussion of these practical consequences, see this Journal, XIII, p. 136, and XXIV, p. 87.³ Cf. W. Brown and W. Stephenson: "A Test of the Theory of Two Factors," Brit. J. Psych., XXXIII, 1953, pp. 352-370; G. Thomson, loc. cit. sup; L. L. Thurstone: Multiple Factor Analysis, 1947, pp. 421f. As both Brown and Thomson indicated, their change of front was partly the effect of the change in physiological views regarding cerebral localization (notably the conclusions of Lashley in regard to 'mass action' to which they both refer, and Head's drastic criticisms of the 'cerebral map-makers'). Thurstone and his followers, on the other hand, seem indifferent to biological, physiological, or experimental evidence, and prefer to rely exclusively on statistical analysis.
evidence for the concept of intelligence

the teacher was apt to lay the blame on what Dr. Ballard dubbed the 'general factor of laziness.' Conversely, when a bright child forged ahead in all he undertook, he found himself applauded as a paragon of industry and held up to his fellows as a model of zeal: 'genius,' said the apostles of the gospel of work, 'is just an infinite capacity for taking pains.'

This interpretation was elaborated in some detail by Maxwell Garnett, Pearson's brilliant assistant, and one of the ablest champions of the doctrine of a general factor. After re-analysing a good deal of the available data, he came to the conclusion that the factor was after all a factor of Will rather than of Intelligence, and affected moral behaviour quite as much as intellectual success. It was largely as a result of his discussions with Garnett that Spearman eventually dropped his earlier interpretations ('sensory discrimination' in his first paper, 'neural plasticity' in the second) and proposed instead a hypothesis of 'mental energy.'

But a re-analysis of existing data, coupled with a priori arguments, could scarcely suffice to settle the question, either one way or the other. Accordingly, in our later experiments, Mr. Moore and I correlated assessments for intellectual performances with assessments for physical, temperamental, and moral qualities. This time most of the cross-correlations were certainly positive, though never very large: it seemed, in fact, as if there was a small but far more comprehensive general factor—a super-factor, as it were—making for excellence in every direction, while the older and more conspicuous factor for cognitive efficiency now appeared simply as a broad group factor, confined to cognitive activities alone: in short, the so-called 'general cognitive factor' turned out to be merely one of the largest of a number of 'group factors' varying in extent and size (2, p. 19). At the same time, another broad group factor emerged underlying the temperamental and moral assessments: this was obviously identifiable with what we had previously called 'the general factor for emotionality.' No sharp division appeared, separating affective characteristics from conative. And the so-called cognitive factor was found to be quite as prominent in tests of practical efficiency as in tests of intellectual activity in the narrower sense.

In the light of this further evidence, Garnett's arguments no longer required us to surrender the idea of a cognitive factor. But it certainly seemed necessary to revise the implications conveyed by the word cognition. The basic contrast seems to lie, not so much between cognitive processes and non-cognitive (i.e., affective or conative) in the old introspective sense of those terms, but rather between the capacity for adapting, guiding, or directing mental activities, by means of discriminative and integrative processes, and the capacity for responding promptly, actively, and energetically. Some such distinction was implicit in Spencer's antithesis between mental mechanism and mental force (or, as the Americans preferred to call it, 'drive'). It was, indeed, the distinction originally laid down by Plato. And, in the absence of more appropriate English names, it is tempting to borrow from the Greek, and speak of a general 'cybernetic' factor and a general 'dynamic' factor.

---

1 J. C. M. Garnett: Proc. Roy. Soc., A, XCVI (1919), pp. 102f. Cf. also id., Education and Citizenship, 1921, pp. 476f. It should be noted that in all his writings Garnett, one of the noblest quakers of his day, invariably placed ethical considerations first.

2 I.e., a factor for guiding or controlling: see above, Sect. I (1). On the basis of purely observational and experimental work with children, Professor Piaget seems to have reached a very similar interpretation of the traditional antithesis between cognitive and affective processes: cf. The Psychology of Intelligence, 1950, pp. 4f.
The evidence we have so far considered seems fully to vindicate the notion of a 'general cognitive factor.' However, during the last fifteen years or so, the most frequent object of attack has been the assertion that this general factor is largely, if not wholly, innate. This line of criticism is partly an after-effect of the doctrines popularized by the behaviourist school, which dominated psychology for so long in the United States. Educational writers in this country still quote Watson's well-known pronouncements: "We no longer believe in inherited capacities . . . All have equal chances at birth." Watson, however, overstated his case. A doctrine of perfect equality in regard to innate mental traits would fly in the face of all biological experience: throughout the animal kingdom, except where the characteristic is absolutely essential to life, innate differences between individuals are the invariable rule.

Twins and Siblings Reared Together and Apart.—In an earlier issue of this Journal I summarized the six or seven converging arguments which can be adduced in support of the inheritance of general ability. The most logical method of investigating such a problem is to keep each of the two variables constant in turn, and compare the results. Let us, therefore, take measurements first for children of identical heredity brought up in different environments and secondly for children of different heredity brought up in the same environment.

In the paper just cited, I gave correlations obtained originally from surveys in the London schools, and supplemented them by further data collected by Miss Conway, who had been responsible for the final computations. Thanks to numerous correspondents, she has since been able to increase the number of cases, particularly for the small but crucial groups of monozygotic twins reared together or apart. The total numbers now amount to 984 siblings, of whom 131 were reared apart; 172 dizygotic or two-egg twins, all reared together; 83 monozygotic or one-egg twins reared together, and 21 reared apart. By way of contrast, she has also secured data for 287 foster children.

1 *Behaviourism* (1931), pp. 99f. Watson goes on to guarantee that "given my own world to bring them up in," he could train any healthy infant to follow any type of profession—"doctor, lawyer, artist, regardless of abilities or ancestors." Without going so far as this, Dr. Blackburn, Dr. Fleming, Dr. Heim, and a large number of sociological writers, appear to accept the general behaviourist view; but it should be noted that even Watson slipped in a few reservations which his more ardent disciples commonly omit.

So far as individual psychology is concerned—apart from the discredited claims of the Iowa school—no new facts have been responsible for this remarkable change of view: it seems rather to be an incidental symptom or consequence of an equally remarkable change in the general climate of opinion. In psychology as in politics, the pendulum of fashion swings to and fro; and the vacillations roughly synchronize. During the nineteenth century, the associationists preached an egalitarian doctrine, and three reform bills were passed. Then the close of the century witnessed a reaction; and we ourselves are witnessing the counter-reaction. An excessive emphasis on heredity has now been succeeded by an equally excessive emphasis on environment. Apparently it is difficult to give due weight simultaneously to each.


3 Of the monozygotic twins, only nineteen were found in London; and, owing to the distances involved, we have been obliged to depend for measurements of the rest either on research-students or on local teachers and doctors (to whom we must extend our sincerest thanks). As a result, the correlations for this group may have been somewhat reduced. There is a natural prejudice against separating twins, especially if their sex is the same; and we should like to repeat our appeal for further cases. Although the handful of monozygotic twins reared apart is decidedly small (and it is the outcome of a quest that has lasted for over forty years), the differences between the correlations for this group and the rest are for the most part statistically significant.

The figures for head-length, head-breadth, and eye-colour are based on much smaller numbers in every batch. Eye-colour (assessed by the methods described in my paper in the *Eugenics Review*, XXXVII, 1946, pp. 149f.) was added because, of all readily observable traits, it is immune from environmental influence.
Evidence for the Concept of Intelligence

The correlations are set out in Table I. Since one or two writers apparently think that the figures obtained by American investigators imply different conclusions from those that I have drawn, I have also included the correlations obtained by Newman, Freeman, and Holzinger (15).

<table>
<thead>
<tr>
<th>Measurement</th>
<th>A—Burt and Conway</th>
<th>B—Newman, Freeman &amp; Holzinger</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mental (Intelligence)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group Test</td>
<td>-944</td>
<td>-922</td>
</tr>
<tr>
<td>Individual Test</td>
<td>-921</td>
<td>-910</td>
</tr>
<tr>
<td>Final Assessment</td>
<td>-925</td>
<td>-908</td>
</tr>
<tr>
<td><strong>Scholastic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Attainments</td>
<td>-898</td>
<td>-955</td>
</tr>
<tr>
<td>Reading and Spelling</td>
<td>-944</td>
<td>-507</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>-862</td>
<td>-883</td>
</tr>
<tr>
<td><strong>Physical</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height</td>
<td>-957</td>
<td>-981</td>
</tr>
<tr>
<td>Weight</td>
<td>-832</td>
<td>-973</td>
</tr>
<tr>
<td>Head Length</td>
<td>-963</td>
<td>-910</td>
</tr>
<tr>
<td>Head Breadth</td>
<td>-978</td>
<td>-908</td>
</tr>
<tr>
<td>Eye Colour</td>
<td>1.000</td>
<td>-880</td>
</tr>
</tbody>
</table>

As regards intelligence, the outstanding feature of the table is the high correlation between the assessments for identical twins even when they have been reared apart: it is almost as high as the correlation between two successive testings for the same individuals. Between non-identical twins the resemblances (at any rate with our own data) are not much closer than those between ordinary brothers and sisters. Nevertheless, environment is not entirely without effect, particularly when the assessments have been obtained by written tests applied

---

1 Dr. Heim, referring to the American inquiry, states that "when young monozygotic twins are separated ... the differences between their scores are as great as those between unseparated dizygotic twins." But it will be seen that, in point of fact, both with the group test (Otis) and with the individual test (Stanford-Binet) the figures there given for the separated monozygotic twins are appreciably higher than those for the unseparated dizygotic twins, even though their figures for the latter are larger than those of most other investigators.

The figures obtained for twins in the most recent and extensive studies of twins carried out in Great Britain seem in the main to agree with our own. Herman and Hogben report with the Otis group test a correlation of 0.66 for twins of like sex and only 0.53 for twins of unlike sex: if we suppose that about half those of like sex were non-identical, this suggests a figure of about 0.80 for the identical twins (12). Maxwell analysed data obtained with group tests for 468 twins during the Scottish Survey, and found correlations of 0.73 for twins of like sex and 0.63 for twins of unlike sex: as he observes, the latter value is "a little higher than that found in most other studies" (19).
to whole groups. The effect is obvious when we compare the correlations for children reared together and children reared apart. And it might be thought that in the correlations obtained from unrelated children reared in the same homes we have a direct indication of its actual amount. In all probability, however, such correlations mainly reflect the method of placement: a dull or defective orphan would not be boarded out with a highly intellectual family.

The figures for physical measurements, at least in our own data, show very similar trends: with the American data the correlations are somewhat higher, but the disparity is seldom large.1

The results obtained for the scholastic tests, both in the American inquiry and in our own, present a striking contrast. In our own inquiry the correlations for siblings and non-identical twins reared together are actually higher than those for the identical twins who have been reared apart. And it may be instructive to note that the correlations which are most conspicuously increased by similarity of home environment are those for verbal or literary attainments; those for arithmetical attainments are, if anything, increased more by similarity of genetic constitution.

Figures like the foregoing provide ample evidence that individual differences in general intelligence are in part at least inherited, and that they are affected by environmental differences much less than are school attainments. However, the mere fact of hereditary influence the more sober critics do not deny. What they question is whether its amount is really large enough to be of any practical consequence either in the sphere of education or in later civic life.

Now I believe that a good deal of the difficulty arises because both the opponents of mental inheritance and its advocates still cling to wholly out-of-date notions of what is to be understood by such a phrase. Terms like heredity and variation, which played such crucial roles in the theories of Darwin, Spencer and the earlier biometricians, continue to be used by modern biologists, but their implications have radically changed. Moreover, the few educationists who appreciate the relevance of this change seem to be quite uncertain how far the newer theories have undermined the older inferences of the Galton-Pearson school.

The Hypothesis of Multifactor Inheritance.—Galton at the very outset of his work noted that in nearly all mental characteristics the observable differences between individuals are differences of degree rather than of kind, and proposed a scale of continuous variation in place of the traditional schemes of discontinuous types. Now, during the first two decades of the century, both the advocates of the new Mendelian hypothesis, and its opponents, originally supposed that the particulate theory of heredity, and the basic principle of segregation, were incompatible with continuous variation in an inheritable trait. Thus, Pearson and the biometric school contended that, even if true, the Mendelian hypothesis must be exceedingly limited in its application, and could have little or no bearing on normal psychology. On the other hand, the earlier Mendelians, De Vries, for example, believed that, since the Mendelian mechanism must underlie all forms of inheritance, no continuous variations could ever be inheritable; and this argument is still adduced by those who reject the inheritability of intelligence, because (so they assume) the very fact that variations in intelligence are continuous shows that they are produced by purely environmental agency.

1 The high correlation for physical measurements obtained by Newman and his colleagues with non-identical twins is a little surprising. Lauterbach’s figures agree more closely with my own. His correlations for twins of like and unlike sex are, for height, 0·80 and 0·53; for weight, 0·89 and 0·80 (Genetics, X, 1925, pp. 525-588).
Evidence for the Concept of Intelligence

Now, in spite of their undoubted importance for genetic and agricultural research, analogies drawn from the study of domesticated animals and plants may be highly misleading when we turn to human genetics. Very naturally, the characters that first caught the eye of the Mendelian experimentalist were qualitative traits, attributable each to some single factor or 'gene' which produces its own visible and distinctive effect. But there is no reason why genes should not exist whose separate manifestations evade our present methods of discrimination-systems of polygenes, and whose effects are small, similar, and cumulative.1 If the number affecting the same trait were large, the result would be that observable variations in that trait would appear continuous, and the frequency-distribution of the measurements would approximate to the normal curve.2 This is fully in keeping with the conclusions reached by Galton and his followers. It may be shown, says Galton, "that the distribution of human qualities and faculties (qualities like height and head-length, faculties like strength, visual acuity, or general ability) is approximately normal" (9, p. 32; 10, pp. 59, 201).

Manifestly it is impossible to check the existence of such genes by direct Mendelian methods; but, with the aid of statistics, we can discover whether the apparent effects are in accordance with Mendelian principles. Suppose, then, that a child's endowment of intelligence is dependent, not on a single pair of genes, but on many such pairs, each segregating in the usual fashion, and all affecting the same observable trait; and suppose too that one member of each pair (designated by a capital letter) would, if present, add a small quantity to the net result, while the other (designated by a small letter) would deduct an equal quantity. Then, for any given individual (or 'phenotype'), the total mount of intelligence would be proportional to the number of capital letters specifying the 'genotype.' Hence, if there were only three pairs of genes, the brightest individual would have a genetic constitution represented by AABBCC, the dullest a constitution represented by aabbcc, and the average person a constitution represented by AaBbCc. Assuming that mating is random and that there is no 'dominance,' the frequency of each genotype could be deduced by expanding the product (A+a)2 (B+b)2 (C+c)2: it would, in fact, be proportional to the binomial coefficients, 1, 6, 15, 20, 15, 6, 1. With n such pairs of genes there would be 2n+1 classes. And, as n increases, the binomial distribution will approach

1 The possibility of multi-factor inheritance was mentioned by Mendel in his discussion of the colouring of white, red, and purple flowering beans. It was first demonstrated by H. Nilsson-Ehle in hybridization experiments on oats and wheat (Kreuzungsuntersuchungen an Hafer und Weizen, 1909); and the cardinal principles were elucidated more fully by E. M. East in studies of the corolla-length in the tobacco-plant ('Size-inheritance in Nicotiana.' Genetics, I, 1915, pp. 164-176). The first to point out the importance of such a theory for human genetics appears to have been C. B. Davenport ('Inheritance of Stature,' Genetics, II, 1917, pp. 313f.). The number of genes which the theorist may legitimately postulate is now known to be far larger than was formerly thought: the banana-fly, Drosophila, is estimated to possess between 5,000 and 10,000; and man may have six times as many.

2 For those who are not familiar with recent work in genetics, a brief explanation may be helpful. H. G. Wells, in one of his short stories, tells how an engaged couple hailing from North Wales—a Mr. Price-Jones and a Miss Evan-Roberts—plume themselves on bearing the family names of both their fathers and their mothers. But, they ask, how are they to christen their prospective children? The minister who is to marry them suggests that each child should take one surname from the male parent and one from the female, and that a coin should be tossed to decide the choice. Now let us apply the same principle to the case where a Mr. Price-Jones had married a Miss Price-Jones: the possible names for the children would be Price-Price, Price-Jones, Jones-Price, and Jones-Jones. This is exactly parallel to the way single genes are transmitted. Put A for Price and a for Jones. Then, when Mr. Aa marries Miss Aa, the possible recombinations are AA, Aa, aa, and aa: since Aa and aA are equivalent, the resulting proportions given by the toss will be 1:2:1.
more and more closely to the normal curve. But this, as we shall see, is only part of the story.

_The Frequency-distribution for the General Population._—Modern critics of the Galtonian view usually start by attacking the theory of normal distribution. Dr. Heim, for example, assures us that it is a sheer assumption, "though not explicitly recognized as an assumption"; quite unwarrantably (she says) it has got "hailed as a scientific discovery, despite the fact that frequency distributions depend mainly on the system of scoring adopted." Mr. Richmond makes much the same point. To ensure this "a priori principle" (he says) the psychometrist "tinkers with the test material"; as a result "measurements are normally distributed, simply because the test has been so constructed that they must be so distributed" (11, 18).

Such arguments betray a singular indifference to the facts. In this country the first attempts to secure objective evidence about the distribution of test measurements were those made during my surveys of London schools. The chi-squared test was applied; and (as I pointed out at the time) the results disclosed quite plainly that such measurements are _not_ distributed in exact conformity with the normal curve. The most conspicuous departure appeared in the lower tail of the curve, where, owing to an excess of dull and defective pupils (by no means invariably of a pathological type), the frequencies were much larger than the expected values.1 When the defectives are omitted, then the resulting curve approximates more nearly to the normal, though the fit is still far from perfect. This _approximate_ normality (which was all that Galton claimed) is thus not 'an a priori assumption' but an empirically demonstrated fact.

On examining the frequency curves for intelligence, therefore, we seem compelled to envisage two kinds of inheritance—unifactor inheritance and multifactor inheritance. If I may repeat what I have said elsewhere, "both the form of the distribution and the correlations obtained are very much what we should theoretically expect were these graded measurements, mainly though not wholly, determined by a very large number of similar genes; while in certain instances and in certain forms (as independent evidence from pedigrees suggests) mental deficiency may occasionally act like a dominant, or, still more frequently, like a recessive, and in some even be sex-linked": in this double mode of transmission, so I suggested, the inheritance of intelligence seems to resemble the inheritance of stature (3, p. 81). Moreover, as with stature so with intelligence, the observable measurements are in some degree modified by non-heritable influences. In the case of stature, the excessive frequency of very short persons is due partly to single genes (as with the achondroplastic dwarf, where the condition is dominant, and the ateleiotic dwarf, where it is apparently recessive), partly to environmental and pathological causes (as with rachitic or undernourished children), and sometimes to both (as with the cretin); and precisely the same types of causation are traceable in the dull and mentally deficient.

---

1 A typical curve is that printed by Mayer Gross, Eliot Slater, and Martin Roth in their recent textbook on _Clinical Psychiatry_ (1954, p. 56): the diagram is reproduced from one of my earlier surveys and based on over 3,000 cases; the irregularities are clearly visible.

Mr. Richmond cites as an example of 'tinkering' the revised version of the Binet-Simon Scale. But the tests were standardized with no reference whatever to normality: the assumption made was that, between the ages of 5 and 12, the annual increments are approximately equal. With properly constructed group tests, the items are selected (often by elaborate scaling techniques, such as paired comparison or its equivalents) so as to increase more or less uniformly in difficulty. Even in mechanical tests like erasing o's and e's in a page of pied print, where there can be no suspicion of 'tinkering with the scale,' the distributions are still approximately normal.
The Frequency-distributions for Parents and Siblings.—The possibility of polygenic determination was not overlooked by the biometric school. Galton himself was convinced that "inheritance may be described as largely, if not wholly 'particulate'" (10, p. 7). And Karl Pearson carried out a theoretical study of the statistical consequences of multifactorial inheritance (16). He concluded, however, that the correlations actually observed both between parents and their offspring and between children and their own brothers or sisters were far too high to be explicable by any such hypothesis. But, as now seems plain, his deductions were partly invalidated by certain untenable assumptions and several undue simplifications. To begin with, he tacitly assumed that dominance would be complete; furthermore, though keenly aware of the facts of assortative mating, he failed to make correct allowance for its influence; and above all, like most of the earlier biometricians, he failed to recognize the clear distinction between the causes of inheritable variation and their observable effects, between the carriers of heredity and the manifestations of heredity, in short, between what is conveniently called the 'genotype' (the hereditary determinants considered as a system typical of certain individuals) and the 'phenotype' (the kind of individual organism eventually produced by the interaction of the genotype with its particular environment); and it is a failure to recognize the same distinction that is largely responsible for the misconceptions and criticisms which the genetical psychologist encounters to-day.

The examination of the bivariate distributions is greatly simplified if we work with grouped frequencies. It is not difficult to show that, if a large number of genes combine, in the manner described above, to determine the measurements for two related members in a random sample of families (e.g., for parents and their children or for children and their sibs), and the measurements are suitably grouped to yield classes instead of continuous variates, then the frequencies to be expected will be similar to those deducible from a single pair of genes, for which the hybrid state \(Aa\) or \(aA\) is intermediate. Such frequencies, of course, can be readily computed by applying the ordinary principles of probability. The detailed values for multifactor inheritance have, in fact, been deduced by Fisher in his classical paper on 'The Correlation between Relatives on the Supposition of Mendelian Inheritance' (7): a non-technical account will be found in (8).

To ascertain how far the actual results for general intelligence conform with those which are required by the multifactor hypothesis, I have collected assessments for a 1,000 pairs of sibs, representing, so far as possible, a random selection of the London school population. At the same time I have endeavoured, though with poorer success, to secure assessments for at least one parent. Since these proved obtainable for only 954 cases, the analysis has to be limited to this smaller number. On the basis of the measurements, the children were divided

---

1 The inquiry was limited to children between the ages of 8 and 13, and was based primarily on verbal and non-verbal tests of intelligence. The actual measurements were transformed into standard scores (i.e., deviations divided by the standard deviation for each age); and these scores in turn were converted to terms of an I.Q. scale with a standard deviation of 15. Thus, the dividing lines for the three groups are approximately I.Q's of 90 and 110. Borderline cases were specially investigated in the light of the teachers' reports, and doubts resolved by individual testing. For the assessments of the parents we relied chiefly on personal interviews; but in doubtful and borderline cases an open or a camouflaged test was employed. The entire set of data on which the following tables are based were derived from four successive surveys carried out with the assistance of Miss Pelling, Mr. Seymour, Miss Richardson, and Miss Howard respectively. The methods adopted were slightly different in each; and the last was the most accurate. But, so far as the grouped frequencies are concerned, the results disclose no significant changes; hence, it seems legitimate to lump the whole series together for purposes of the present analysis.
into three groups—bright, average, and dull—in the proportions 1:2:1; and a similar classification was adopted for the parents. The percentages we should expect for the bivariate distribution, based on the triple assumption of random mating, Mendelian segregation, and no tendency to dominance, are shown below in Tables IIA and IIIA. They are, it will be noted, in the proportions 1, 1, 0; 1, 2, 1; 0, 1, 1 for parent and child, and 9, 6, 1; 6, 20, 6; 1, 6, 9 for pairs of sibs. On calculating the product-moment correlation for each hypothetical table, the value will be found to be exactly 0·500.

The observed frequencies, also reduced to percentages, are shown in Tables IIB and IIIB: (the perfect symmetry of the latter results from the procedure regularly followed in constructing a table for an intra-class correlation). It will be seen that the observed proportions agree tolerably well with the hypothetical; and, as we shall learn in a moment, the divergences themselves are very much what we should anticipate. The actual correlations, computed from the original data, were, for parent and child, 0·481, and for sibs 0·507 (computed from the pooled frequencies tabulated below, the values would be slightly different owing to the 'coarse grouping').

**TABLE II**

**Bivariate Distributions for Parents and Their Children.**

<table>
<thead>
<tr>
<th>Parents</th>
<th>A.—Theoretical Frequencies</th>
<th>B.—Observed Frequencies.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bright</td>
<td>Children</td>
<td>Bright Average Dull</td>
</tr>
<tr>
<td>Bright</td>
<td>12·5 12·5 0·0</td>
<td>25·0</td>
</tr>
<tr>
<td>Average</td>
<td>12·5 25·0 12·5</td>
<td>50·0</td>
</tr>
<tr>
<td>Dull</td>
<td>0·0 12·5 12·5</td>
<td>25·0</td>
</tr>
<tr>
<td>Total</td>
<td>25·0 50·0 25·0</td>
<td>100·0</td>
</tr>
</tbody>
</table>

**TABLE III**

**Bivariate Distributions for Siblings.**

<table>
<thead>
<tr>
<th>Children</th>
<th>A.—Theoretical Frequencies.</th>
<th>B.—Observed Frequencies.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bright</td>
<td>Children</td>
<td>Bright Average Dull</td>
</tr>
<tr>
<td>Bright</td>
<td>14·1 9·4 1·5</td>
<td>25·0</td>
</tr>
<tr>
<td>Average</td>
<td>9·4 31·2 9·4</td>
<td>50·0</td>
</tr>
<tr>
<td>Dull</td>
<td>1·5 9·4 14·1</td>
<td>25·0</td>
</tr>
<tr>
<td>Total</td>
<td>25·0 50·0 25·0</td>
<td>100·0</td>
</tr>
</tbody>
</table>

A perfect agreement between the observed frequencies and the theoretical cannot possibly be expected, since there must be numerous unavoidable influences, tending partly to increase and partly to diminish the apparent correlation. (i) To begin with, like all mental measurements, assessments for intelligence, however scrupulously checked and adjusted, are in some degree distorted by the unreliability of the methods available. The best estimate for the reliability coefficient is 0·916. If we apply the usual correction for unreliability, the observed values would be raised to 0·525 and 0·554 respectively. (ii) But, as we have seen, the most punctilious attempts to assess 'innate ability'
Evidence for the Concept of Intelligence

(itself a purely hypothetical quantity) cannot entirely escape the effects of different environmental conditions; and, of course, for members of the same family the effects must generally tend in the same direction. How far this may have augmented the apparent correlation it would be hard to say: but the increase must almost certainly have been smaller than the decrease due to unreliability.

(iii) The most elusive tendencies to allow for are those of dominance and assortative mating. Were dominance complete, the expected correlations would be altered to $\frac{q}{1+q}$ and $\frac{1+3g}{4(1+g)}$ respectively, where $p^2$, $2pq$, and $q^2$ denote the proportions of pure dominants, mixed dominants, and pure recessives respectively, and $q+p=1$. Thus, the effect of dominance is once again to lower the apparent correlations; but, unlike that of unreliability, it lowers them by widely differing amounts. Now the initial classification we have adopted makes $p=q=\frac{1}{2}$. Substituting this value in the fractions given above, we obtain, for the expected correlation between parents and their children, a coefficient of $\frac{1}{3}$, that is $0.333$, and for the expected correlation between children and their brothers or sisters a coefficient of $\frac{5}{12}$, that is $0.416$. The observed values are significantly higher.

(iv) What then can have raised the absolute values to this high level? The most likely answer is assortative mating. How then can its presence be verified and its influence assessed? One of its calculable results would be to increase the variance of the younger generation. Now, if we may trust our rather crude measurements, the variance of the parents is only $12.3$ I.Q., whereas (in virtue of the mode of standardization) the variance of the children is $15$ I.Q. This is tantamount to an increase in the filial generation of about $22$ per cent. Spouses, it appears, prefer partners whose intelligence in some degree resembles their own. The actual amount of the resemblance can be estimated by calculating the correlation between husband and wife. In the earlier surveys it was well over $0.40$; in the later somewhat below.

Now, as we have seen, dominance, like unreliability, tends to reduce the correlations between parent and child and between one sib and another; but, unlike unreliability, it reduces them to different extents. On the other hand, the effect of assortative mating, like that of similar environment, is to increase the correlations, and to increase them by amounts that make them more nearly equal. The net result can be estimated, if we treat the contributory variances as additive components, and then apply the ordinary principles of factor analysis. A little calculation indicates that the ultimate effect of assortative mating would be to add a small amount to both correlations, viz., in the present

---

1 The critic usually supposes that intelligence-tests are considerably affected by cultural differences in the testees' environment. But, if the tests have been properly constructed and their pronouncements properly checked and adjusted, such effects are almost negligible. The influence of unhygienic conditions in early infancy is a more likely source of error, for which it is difficult to allow. The statement in the text was based on indirect attempts to estimate the upper limit for environmental influences by methods which need not be detailed here.

2 These are the theoretical values deduced by Pearson in the paper already cited (16). He rejected them, and with them the assumptions on which they were based, because they fall far below the correlations he had empirically obtained for numerous traits showing continuous variation. Yule, however, pointed out that if, instead of postulating complete dominance (as Pearson had tacitly done), we postulate complete absence of dominance, both the theoretical values would be raised to $0.500$; and this would accord far better with Pearson's own figures (Report of Conference on Genetics, 1906). Yule's assumptions make the relevant conditions even simpler than Pearson's; but in view of recent results, it seems pretty certain that they are far more complex.
instance about 0·15 and 0·09. With complete dominance, this would raise the theoretical values from 0·333 and 0·416 to about 0·48 and 0·52 respectively. And these figures tally reasonably well with those observed.¹

The Relative Importance of Heredity and Environment.—We now reach our final problem: what proportion of the total variance shown by the children is attributable to genetic conditions as contrasted with environmental? In recent discussions on this point, two important considerations are frequently ignored.

(i) If the observed correlation between parent and child is 0·481, we might infer that each parent contributes 0·481² = 23 per cent. to the total variance. And if the mating were random, the two parents together would contribute 2 × 23 = 46 per cent. But since, as we have seen, there is a correlation of at least 0·40 between fathers and mothers, part of the influence of one parent must overlap with that of the other, and consequently should not be included twice. Making due allowance for the overlap, we may estimate the contribution deducible from the assessments for the two parents as about 45 per cent. at most. Now it is often inferred that the remainder of the variance must, therefore, be ascribed to non-inheritable factors, that is, to the influence of the environment. But with the mode of transmission we have assumed, not only the parents but also the grandparents and remoter ancestry must contribute something to the variance. A simple algebraic deduction from the postulates of multifactorial inheritance will show that the total effect of parentage and ancestry may be directly measured by the correlation between sibs. The observed correlation, it will be remembered, was 0·507. According to our findings, therefore, about 51 per cent. of the variance must be contributed by such factors.

(ii) But, even so, it would be quite mistaken to assign the whole of the residue (49 per cent.) to environmental influences. By an odd paradox, not only the similarity between siblings, but also their differences are largely the outcome of their genetic constitution. Thus, arguing from Mendelian principles, we should definitely anticipate a frequent lack of resemblance between one sib and another owing to the segregation of those factors in respect of which the parents are heterozygous. After computing a rough estimate for this additional contribution, I calculate that in all at least 75 per cent. of the entire variance must be due to genetic influences, probably far more.²

It must be frankly owned that, with a sample covering under a thousand cases, the somewhat speculative balancing of the accessory factors that affect the correlations here obtained can make no pretence to be either accurate or

¹ It might be suggested that the resemblance between brothers and sisters appears greater than that between parents and their children, because children of the same family are brought up together and may even go to the same school. Similarity of schooling might no doubt affect the correlations for cultural and educational tests, as indeed the figures in Table I suggest; but (for the reasons given above) it cannot appreciably affect the results obtained for intelligence. Certainly the assessments for siblings who have gone to the same schools reveal no higher correlation than the assessments for those who have gone to different schools.

² The theoretical considerations on which such calculations should be based are clearly set out by Fisher in the paper already cited (7). It may be observed that the figure for the residual contribution which we have thus reached, namely 25 per cent., would imply a correlation with the conditions causing it amounting to \(\sqrt{0·25} = 0·50\). But a direct calculation of the partial correlation between favourable environmental conditions and the assessments for intelligence proves to be well under this figure. Hence, the final figure reached above for genetic influence leans definitely to the conservative side. Fisher's formulae would subdivide the contributions to the total variance into (i) genotypes 49 per cent., (ii) dominance 28 per cent., (iii) assortative mating 19 per cent., leaving for (iv) environment only 4 per cent.
conclusive. My aim has rather been to adumbrate a line of reasoning that merits closer consideration and further research. But, even as it stands, the analysis I have made, supplemented by the other evidence that I have mentioned, seems to me to afford a strong corroboration for the view I have indicated, namely, that human intelligence, like human stature, is determined largely though not wholly by multifactorial inheritance.1

Conclusion.—I have now reviewed the wide variety of evidence—observational, introspective, and experimental, biological, physiological, and statistical, bearing on our initial question. The results are mutually supporting; and, apart from certain minor modifications or extensions, seem abundantly to confirm the threefold hypothesis that I tentatively put forward over forty years ago in the forerunner of this journal2: namely, that there is a general factor making for efficiency in all mental activities, that this factor is essentially cognitive or directive, and that the greater part of the individual variance found in this factor is attributable to differences in genetic constitution. This triple conclusion suggested a modernized formula for the abstract conception to which so many different writers had been led, viz., ‘innate, general, cognitive ability.’ If, therefore, we are to retain the word ‘intelligence’ as a technical term in psychology, this still seems the best definition.

III.—Summary.

The main steps in the argument may be epitomized as follows:

1.—Evidence from different branches of psychology leads to the notion of a mental capacity that is (i) cognitive, (ii) general, (iii) innate.

2.—Each of these three characteristics has been amply verified by statistical research.

3.—As the history of the word shows, intelligence was a technical term put forward to designate a technical concept: and the meaning given it, implicitly or explicitly, by leading authorities from Cicero and the scholastics to Spencer, Galton, and Binet, suggests that it furnishes the most convenient name for the concept thus reached.

1 As I suggested in my earlier paper, it is urgently desirable that similar methods should be employed to investigate the presence of a general cognitive factor in lowlier animals, and, if possible, to determine its mode of inheritance. The few researches so far carried out point to conclusions similar to those reached above. R. L. Thorndike has calculated correlations between tests of learning, strength of drive, etc., in albino rats, and finds a general factor of learning ability and two supplementary factors (Genet. Psych. Monogr., XVII, 1935, pp. 1-70). Vaughan (Comp. Psych. Monogr., XIV, 1937, pp. 1-41) and Tolman and others (ibid., XVII, 1941, pp. 1-20) have also published tables of correlations for the performances of rats: their figures are fully consistent with the theory of a general factor, though the investigators themselves prefer an analysis in terms of overlapping group factors.

To test the hypothesis of multifactorial inheritance, Tryon has carried out experiments on maze-learning, which he regards as a test of general ability, and has repeated them with successive generations. He attempted first to secure two strains, of bright and dull rats respectively, by selective inbreeding. After seven selections and seven generations, he found practically no overlapping between the distribution-curves for bright and for dull. He then, as it were, reversed the procedure, crossing the two strains, and testing two further generations. It is true that the variance exhibited by the F1 generation seemed too great to be explained wholly by non-genetic influences, and much greater than would be expected had the method of inbreeding been successful. Yet, on the whole, as he contends, the results seem to support the multifactorial hypothesis (R. C. Tryon, “Genetic Differences in Maze-Learning in Rats,” Thirty-ninth Yearbook of Nat. Soc. Study of Education, 1940, pp. 113f.; cf. also E. G. Brody, “Genetic Basis of Spontaneous Activity in the Albino Rat,” Com. Psych. Monogr., XVII, 1942, No. 5.).

2 See J. Exp. Pedag., I, 1911, pp. 93f.: Cf. also 1.
4.—Apart from comparatively rare and abnormal variations, differences in intelligence as thus defined seem to depend on the combined action of numerous genes whose influence is similar, small, and cumulative—a hypothesis that is fully borne out by the frequency-distributions obtained for parents, siblings, and the population as a whole. And on this hypothesis not only the similarities between relatives but also their dissimilarities will be largely due to genetic factors.

5.—It is essential to distinguish between intelligence as an abstract component of the individual’s genetic constitution (g) and intelligence as an observable and empirically measurable trait (G). The evidence indicates that at least 75 per cent. of the measurable variance (based on carefully checked assessments) is attributable to differences in genetic constitution, and less than 25 per cent. to environmental conditions.

IV.—References.