UNIVERSITY OF EDINBURGH.

AN INVESTIGATION INTO TEMPERAMENT: ITS RELATION TO INTELLIGENCE, AND SPECIAL ABILITIES.

BY

FLORENCE E. MACNEILL, M.A., B.ED.

THESIS PRESENTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

JUNE 1942.
ACKNOWLEDGMENTS.

The writer wishes to express her thanks to Dr Mary Collins for lending her the puzzle on the basis of which the test F.C.C. was devised.

Facilities for testing were also kindly granted by Dr J. Mackie, Rector of Leith Academy, and by Mr Geo. Mowat, Headmaster of Moray House School.

Finally, the writer must especially thank Mr D.N. Lawley for his invaluable assistance with the statistical computations.
INTRODUCTION: Purpose of Investigation.
Definition of Temperament.

Ch. I. SURVEY OF THE DIFFERENT APPROACHES TO TEMPERAMENT.
      pp. 5-56
      a) Descriptive Typologies (mainly German)
      b) The Isolation of Traits and the Statistical
          Approach (English & American)
      c) Current Methods of Assessment.
          1) Measurement of Traits.
          2) Verbal-Associative process.
          3) Intelligence test scores & Behaviour
             during tests.
          4) Analysis of Psycho-Motor factors.
          5) The Rorschach technique.
          6) Measuring 'single levels'.
Ch. IV.  THE INTER-ACTION WITH TEMPERAMENT OF INTELLIGENCE, AND SPECIAL ABILITIES.

Ch. V.  DESCRIPTION OF TESTS USED IN THE INVESTIGATION.

a) A Battery of Performance Tests.
b) Group Tests used in the Investigation.
c) Value of Performance Test Situation.

Ch. VI.  GENERAL RESULTS WITH COMPLETE TEST BATTERY.

a) The Group tested.
b) Test distributions.
c) Statistical Results
   1) Factor Analysis.
   2) Application of Method of Maximum Likelihood to test significance of factor loadings.
d) Sex differences.
e) What are the tests testing?

Ch. VII.  METHOD OF TEMPERAMENT ASSESSMENT.

a) Description of traits.
b) Special reactions.

Ch. VIII.  ESTIMATION OF A GENERAL TENDENCY OF TEMPERAMENT.

a) Comparison with teachers' estimates.
b) Calculation of Significance.
c) Conclusion.

Ch. IX.  SOME /
Ch. IX. SOME INDIVIDUAL RESULTS.

Ch. X. CONCLUSIONS AND SUMMARY.

In the autumn of 1928 the writer decided that she would

...
INTRODUCTION.

In the autumn of 1939 the writer decided that she would like to attempt a piece of research into the nature of temperament, for it seemed to her then that there was a pressing need in Applied Psychology and Education for some such estimate to supplement the intelligence test, and if possible, for one which would be more objective than those hitherto in use.

Now that we are fighting to retain and develop what is of value within our culture, it seems more than ever necessary that man should "know himself" and that as much effort as possible should be exerted to fill up the enormous gap between scientific achievement on one hand, and man's behaviour as a social being on the other. Clearly, speculation about the non-rational and instinctive factors in an individual's make-up is not enough; in so far as it may be possible to pin down any such factors, they will have to be subject to control and prediction.

The writer therefore concluded that whatever small contribution she could make to this field would have to be based on the results of direct observation, not on intuition, and as little as possible on pre-conceived theories.

First of all it is necessary to discuss the sense in which we are using the term Temperament. In the narrowest sense it is used to describe excitability or a prevailing mood/
mood, while in its widest application it is substituted for Personality. In this latter connotation too much is encompassed, although of course there is a sense in which all aspects of human behaviour are in some degree, influenced by Temperament. We would prefer, however to think of Personality as the complete psycho-physical make-up of the human being, with emphasis also on its "wholeness", and perhaps best to be described, following Stern, as a "unitas multiplex". This would include all the psychological factors both inherent and acquired which underlie the behaviour of an individual.

This inter-changeability of the terms Personality and Temperament has arisen because of the difficulty of separating levels of origin of behaviour. Some psychologists prefer to investigate certain characteristics, behaviour traits, or "attitudes" in terms of the general personality, without attempting to define their level of origin. This is understandable, for our tests of intelligence, temperament, character, and emotionality are measures of behaviour, and behaviour is a function of the personality as a whole.

Many current psychological definitions of temperament stress its physical basis. McDougall (1) has defined temperament as "the sum of the effects upon mental life of the/

* We shall not discuss the acquired factors which constitute "character", for our psychological analysis takes no account of ethical values.

(1) An Outline of Psychology. 5th Ed. London.
the metabolic or chemical changes that are constantly going on in all the tissues of the body." Combine this with what he calls "Temper", that is, "the expression of the way in which the conative impulses work within an individual" and we have a reasonable starting point. There are three principal ways in which this "working of the impulses" varies from man to man, namely in respect of (1) strength, intensity, or urgency (2) persistency and (3) affectability. Affectability indicates the degree to which impulses are influenced by pleasure or pain, and seems to contain some emotional element. Volitional behavior is also covered by these aspects of Temper, for McDougall refers to steadfastness and persistence, irrespective of motives, as resulting from different "Temper" patterns.

Rather better than this, however, is Kretschmer's definition, "Temperament denotes the general affective disposition characterizing the entire individuality, the special selective affective levels on which the individuality reacts for preference in conjunction with the neuro-humoral substrata". Kretschmer thus stresses both the affective nature of temperament, and its neural and endocrine basis.

Advances in endocrinology ought, as time goes on, to help us within our field, for endocrine secretions are the most probable organic equivalents of instincts, to which of course reference must always be made when the question of

of emotion arises. And no theory of Temperament seems possible without taking emotionality into account.

We can see, therefore, how incomplete is our knowledge in the temperamental field, and because of this the modern conception of Temperament is in the main heuristic, and we cannot set too rigid a standard in deciding what phenomena are to come under the heading of Temperament.

On the other hand, to the present investigator, at least, it seems artificial to attempt to examine the conative aspect by itself, for cognitive factors such as perseveration become involved and it is difficult to draw any sort of dividing line. Incidentally, since the writer means the actual activity involved, the term orexis will be used, rather than conation.

The present writer has sought to investigate Temperament in the light of both oretic and cognitive factors. Fundamentally it would appear that an individual's Temperament is based on instinctive drives, later to be environmentally conditioned. One may not necessarily feel compelled to differentiate those instincts, for a descriptive scheme of Temperament can be evolved without a detailed instinct differentiation. It is useful, however, to start off with some sort of cumulative concept of instinctive energy such as that of (1) Burt: "Human life must not be regarded as a composite of clean-cut blocks; ... life is a single, flowing stream/"

stream, not an aggregate of discrete faculties or unrelated reflexes. The several emotions themselves are but specialized differentiations of a primaeval emotional energy - a fundamental will to live." Instinct is thus thought of in a cumulative sense, as a "central core of impulse", as Ginsberg puts it. This does not necessarily mean that we are assuming "General emotionality" in any other sense than as an inner instinctive 'potential', which may be tapped in a specific overt action, or actions.

CHAPTER I. SURVEY OF THE DIFFERENT APPROACHES TO TEMPERAMENT.

It is important that we should make a short survey of the more notable work which has been done on Temperament, in order to see whether, amongst the many different aspects selected and given prominence, there are any concepts which are common to most of the theories, or at least to a group of theories. It might then be feasible to assume that some sort of temperamental "reality" might exist amongst those concepts, and it would be interesting to compare those with the results of our own experimental investigation.

There have been theories of Temperament since classical times and the temperamental differences have usually been considered to be of a qualitative rather than of a quantitative order. The idea that humanity can be divided on a rough basis of two inherently opposed types, each with its/
its characteristic temperamental bias, is as old as the concept of Temperament itself. We find its modern expression in Dr. James' description of the contrasting types of Will, which he labelled the "explosive" and the "obstructed" types. The Explosive will be subdivided under two headings. One of these is that of defective inhibition, in which impulses seem to discharge so promptly into movements that inhibitions have no time to arise. The other is exaggerated impulsion, and this James describes as follows: "Disorderly and impulsive conduct may come about where the neural tissues preserve their proper inward tone, and where the inhibitory power is normal or even unusually great. In such cases the strength of the impulsive idea is preternaturally exalted, and what would be for most people the passing suggestion of a possibility becomes a gnawing, craving urgency to act."

This broad division made by (1) James is very relevant to our own investigation, for we shall see that it bears marked similarity to what we observed in the psycho-motor field. His scheme of course, has only some descriptive value; it has no explanatory significance.

Baldwin made a broad division between motor and sensory individuals. The motor type tends to act promptly, quickly, unreflectively; generally such a person, child or adult, is said to jump at conclusions. Psychologically such a person is dominated by Habit. He is domineering and self-assertive/

(1) Principles of Psychology. Vol. II.
self-assertive - the man of action. "The sensory child is passive, more troubled by physical inertia, more contemplative when a little older, less apt in learning to act out new movements, less quick at taking a hint." The sensory individual is the observer, the thinker.

Very much later, Davenport, in his analysis of Temperament, emphasised as fundamental to its varieties the contrasting trends called by him (after Southard) "hyperkinesis" and "hypokinesis". Hyperkinesis involves increased activity and elevated emotional tone; hypokinesis, decreased activity and lowered emotional tone. Alternating cycles from hyperactivity to depression are not uncommon.

We thus see that descriptive psychology has often cast about for terms applicable to big patterns of organization, comprehensive enough to include many phases in its scope. This is evident in the "introvert" and "extravert" categories of C.G. Jung, and the "schizothymes" and "cyclothymes" of E. Bleuler and E. Kretschmer.

At this point, how, it will be useful to make two broad divisions of the modern work on Temperament into

(1) Descriptive Typology. (Mainly European, particularly German) and (2) The Isolation and measurement of single traits; and the statistical approach. (Mostly confined to England and America).

(1) Inheritance of Temperament.
(2) Körperbau und Charakter. Berlin 1922.
Descriptive Typology.

The inventors of typologies have recognised the heuristic connotations of many of the present temperamental concepts, and studying the subject largely from the neuro-psychiatric angle, have tried to depict the broad lines of the personality as a whole, describing those individuals who have certain outstanding features in common as "types". Kretschmer, as a result of much psychiatric experience described two main types of personality, distinguished from one another on a temperamental basis. His evidence is largely descriptive and he has few figures to support his conclusions. The 2 main types - schizothyme and cyclothyme are held to be connected with certain types of physical constitution, individuals of 'leptosomatic' or long-bodied build, as well as those having certain developmental anomalies (dysplasias) who are predominantly schizothymic, while those of round-bodied or 'pyknosomatic' build are predominantly cyclothymic. Schizothymic temperaments vary principally in their sensitivity (psychaesthesia) ranging from hypersensitivity and excitability to frigidity and anaesthesia. Cyclothymic temperaments, on the other hand vary principally in mood (diathesis) ranging from exaltation to depression.

The schizothymes are in less direct contact with their environment than the cyclothymes. Inner elaborations of external stimuli are more frequently set up, and emotionally-toned/
emotionally-toned systems formed. Hence the reactions of schizothymes are less predictable in terms of the nature of the immediate stimulus than the cyclothymes, who tend to respond more readily and to form few inner elaborations or "complexes". What may be an even more fundamental difference is the peculiar "jerkiness" of all psycho-neural processes in the schizothymes, as compared to their smoothness in the cyclothyme. The variations in the emotional state or sensitivity of the schizothyme are sudden and abrupt.

With the schizothymes although outwardly calm and insensitive for long periods, a sensitive "complex", that is, a psychological pattern having a high affective "charge" for the individual may suddenly cause an abrupt and violent reaction, which may last, in characteristic cases, until a further stimulus seems to "jerk" the individual over to a different reaction. In the cyclothyme, however, the swings of mood, however extreme, take a smoother and more gradually progressive course.

This jerkiness of the schizothyme is not confined to emotional and intellectual processes; it appears also and most conspicuously at the psycho-motor level. The movements of the schizothymes are more abrupt, while the cyclothymes flow more evenly. This is important in assessing temperament since the relatively simple psycho-motor level is more accessible to measurement than the more complex and variable higher/
higher emotional and intellectual processes.

We shall return to these differences in psychic tempo and psycho motility later on, when considering the psycho-motor factors in the behaviour of our own subjects.

Kretschmer finally presents 6 chief temperamental types. These are the hypomanic, who are cheerful, lively and mobile; the syntonic who are the practical realists and tend to be easy-going; the phlegmatic; the hyperaesthetic who react in a touchy, nervous fashion, and tend to be 'delicate' minded and introspective; the schizothymes at the mid-position of the psychaesthetic scale, who are coldly energetic or systematic and logical, or "calmly aristocratic". Finally, we have the anaesthetic who may be cold and frigidly nervous, odd and eccentric, indolent, and affectively weak.

So far Kretschner's views have not been confirmed, except descriptively. A large sample of the 'normal' population would have to be investigated in order to obtain any kind of statistical confirmation. Some confirmation has come, however, from F.J. Wertheimer and F.E. Hesketh in a study of mental patients, and from G.J. Mohr and H.H. Grundlach with convicts.

Usually, in medical and psychiatric work the majority of normal subjects are found to be of mixed types, showing some features of each of the antithetic types in varying degrees. /

(1) "Significance of the physical constitution in mental disease". Medicine, 1926, V, 375-461.
degrees. This, of course, does not invalidate the typological interpretation of personality. We find that Kretschmer himself declares that neither the physical nor the temperamental types are to be regarded as disparate entities, but rather as extreme points on a scale of distribution of complex, genetically determined characteristics, pure examples of any one type being unusual.

(1) Rosanoff also contributed a typology based on psychiatric experience. He attempted to relate varieties of normal personality to psychotic or insane types. He held that normal personalities differ from psychopathic ones quantitatively rather than qualitatively and listed 4 types of personality which might develop into abnormal ones.

The anti-social personality is his first type. It underlies hysterical manifestations and shows itself in social life in lying and swindling and malingering. This may develop into the dual personality.

The cyclothymic personality describes a highly unstable type with ups and downs of four varieties. These are (a) the manic make-up, which includes "brilliant but unevenly endowed individuals with artistic inclinations, nimbleness of spirit and versatility, multiplicity of ideas, alert open-mindedness and spirit of enterprise, artistic ability, kind-heartedness/

kind-heartedness, gay sunny disposition". In the mental hospital this is the exalted type with grand notions. (b) The depressive make-up exists from youth showing "a special sensitivity to the cares, troubles and disappointments of life. These individuals take all things hard and feel the little unpleasantnesses in every occurrence much more strongly than the elevating and gratifying aspects of unconcerned and happy enjoyment, of unreserved surrender to the future". In its exaggerated form this is the "melancholic" sunk mentally and physically in the deepest despair. (c) The irascible make-up exhibits fluctuating emotional equilibrium to an extraordinary degree. This type is hot-headed, easily offended, given to violent scenes with rapid changes in emotional colouring. In the extreme form this is the popular "maniac" type. (d) The unstable make-up is characterized by emotional instability, with a "swing back and forth between the two opposite poles of emotion, now 'shouting with joy to heaven', now 'grieved to death'." The mental hospital extreme of this type runs the gamut from mania to depression.

The autistic or shut-in personality is the third main type in Rosanoff's scheme. This forms the basis of dementia praecox, with mental deterioration, emotional apathy, and dissociative phenomena. This is also the vague, dreamy detached person who shows reduction of external interests and a tendency to lose contact with reality.

The epileptic personality is the fourth category and this type tends to be violent, headstrong, subject to violent likes/
likes and dislikes. There may be brief periods of inspiration or moods of ecstasy or an avalanche of ideas (Dostoievsky, from his own experience, depicts this sort of thing.). Seizures, lapses of memory, and a tendency towards violent acts are also characteristic.

Rosanoff explains that all 4 main categories have value in their normal form. An anti-social trend underlies much of our prudence, diplomacy, and success in commerce he says, while the sensitiveness and expressiveness of the cyclothymic, on the other hand, find an outlet in literature and dramatic art. The autistic personality may shut away every diverting influence and may accomplish great things in scientific research. Finally the epileptic, stubborn and inspired, may become the great reformer or religious leader.

Mixed types are not lost sight of in Rosanoff's scheme, for he states that several trends may unite in one individual. "Mixed types are the rule, pure types the exception", he says.

Essentially what differentiates the normal from the abnormal types are such balancing factors as power of inhibition, emotional control, what he calls durability of mind, and nervous stability.

Again, in Rosanoff's typology, we have a useful descriptive scheme, obviously of value in psychiatric and clinical work, but for more extensive use something more exact is necessary.

We have already mentioned the usefulness of endocrinology in assessing Temperament. Berman has based a typology on/

(1) The Glands Regulating Personality.
on glandular differences. An excess secretion of the Anterior pituitary, for instance, produces the masculine, masterful, and domineering type, while deficiency would cause lethargy and indolence. The Posterior pituitary is responsible for the feminine, imaginative, artistic type, and its deficiency produces coarseness and lack of imagination. The Thyroid gland controls energy, and excess secretion leads to restlessness, lack of control and excessive emotionality. Its deficiency gives the inert individual and in an extreme degree, feeblemindedness. The gland of childhood is the Thymus. If it persists in the adult, infantile qualities persist, such as homosexuality. The Sex glands govern maleness and femaleness, both physical and mental, while the Adrenals are responsible for endurance and vigour. Deficiency produces lack of vitality and neurasthenia.

Again this adds to our knowledge in the field of Temperament, since it is based on experience of such case-material. Obviously the influence of hormones is of outstanding importance, but only, it seems, to the present writer, when linked up with other physiological and psychological knowledge. Obviously one cannot build up an entire psychological theory of Temperament on this foundation alone. Bio-chemical processes are certainly of great consequence from the causative point of view, but taken by themselves alone, they constitute as one-sided a theory as one, for example, based solely on neural processes.

Heymans/
Heymans, Wiersma and the Groningen school have contributed a typology which, in the present writer's estimation, has not received as wide a measure of attention as it merits.

It is based on three variables, Emotionality, Volitional Activity, and Primary or Secondary Function. Secondary Function, the term first used by Otto Gross, is of course the same as the "perseverative tendency" originally described by G.E. Müller.

Heymans says -

(a) In general we call some one Emotionally on the basis of the frequency and strength of his affective reactions, in proportion to their causes.

(b) Active on the basis of the frequency and energy of his activities, in proportion to their motives.

(c) Primary or Secondary Functioning according to the degree in which cognitive and affective processes "perseverate" in proportion to their importance.

The material of the research was obtained from the questionnaire replies of 458 reliable doctors in Holland, who in all gave over 2500 temperamental and personality ratings of individuals from families thoroughly well-known to them.

Having determined for each individual in his group whether he is Active or Inactive, Emotional or Non-Emotional, Secondary or/

(1) Perseveration and its nature will be discussed later on.

or Primary Functioning, 8 groups are sorted out.

(1) Secondary Function  Active  Emotional
(2) "  "  "  Non-Emotional
(3) Primary Function  "  Emotional
(4) "  "  "  Non-Emotional
(5) Secondary Function  Inactive  Emotional
(6) "  "  "  Non-Emotional
(7) Primary Function  "  Emotional
(8) "  "  "  Non-Emotional

There are thus 8 possible combinations of the 3 variables.
Heymans calculates the percentage incidence of every quality in the whole questionnaire, finds the average for the whole group, hence determining the extent to which any one quality is present above or below the average in each.

To the 8 combinations he applied a type name taken from everyday speech or from literature on the subject. These names are therefore merely generic titles, used more for convenience than anything else. They do not denote a unitary quality which persons in each group possess. They indicate the general tenor of the dynamic unity of the qualities, correlated with any one of the 8 variable combinations. The following are the 'types' with the appropriate variables.

Passionate  Secondary Function; Active; Emotional.
Phlegmatic  Secondary Function; Active; Non-Emotional.
Choleric  Primary Function; Active; Emotional.
Sanguine  Primary Function; Active; Non-Emotional.
Sentimental  Secondary Function; Inactive; Emotional.
Apathetic
Apathetic  Secondary Function; Inactive; Non-Emotional.
Nervous  Primary Function; Inactive; Emotional.
Amorphous Primary Function; Inactive; Non-Emotional.

This classification has one outstanding virtue, namely its lack of rigidity. The fact that the 3 variables may come together in any given proportion lends an essentially dynamic aspect to the typology. At the focal point where we would expect the largest number of cases, we find that there, the behaviour qualities of Heymans' cases are most mixed and undifferentiated. This tallies with the common finding that most people are of mixed temperament, which is the stumbling block of most typologies, unable as they are to account for the slight differentiations of the mass.

(1) S. Biesheuvel drew up a questionnaire based on the one used by Heymans with combinations of the 8 variables. He found that those variables did emerge from his data. This of course does not prove the reality of the variables, since if the questions are constructed with the variables in mind, it follows that what is put into the questionnaire will of necessity come out of it. But what is noteworthy is that independent estimates of Biesheuvel's subjects agree to a large extent with his own description on the Heymans basis, and the general conclusion from Biesheuvel's research is that the typology does work. And that, after all, is what is important. A group of phenomena may be divided from very different/

different points of view. It is not a case of one being true, and the others being untrue. What really matters is whether or not any or all are useful in furthering the psychological purpose. It would appear, therefore, that the Heymans classification does classify to a useful extent.

The distinction between an attitude and a trait is indefinite especially when the object towards which the attitude is directed becomes highly generalized. Some traits, however, approach very closely to the category of attitudes. They are modes of adjustment through which the person orients himself to some particular aspect of his life-situation. Extraversion, for instance, as well as being a trait, is at the same time a generalized attitude towards reality.

Jung, the originator of the terms Extraversion and Introversion, has undoubtedly aroused more interest through them, than has any other psychologist, with any other traits during the last 20 years. Many times criticised, they nevertheless continue in current psychological usage, and have found their way also into the layman's vocabulary. This dichotomy has taken the place of the former "objective mindedness" and "subjective mindedness" and of James' concepts of "tough mindedness" and "tender mindedness". It is fairly obvious that Jung's terms have won their popularity because of their stylish compactness and immediately apparent etymology. A man is extraverted "when he gives his fundamental interest to the outer or objective world, and attributes an all important/
important and essential value to it; he is introverted, when the objective world suffers a sort of depreciation, or want of consideration, for the sake of the exaltation of the individual himself. An extravert prefers to participate in the objective real world and in practical affairs, while the introvert has preference for the imaginal world which is rich and creative. The affective life of the extravert is not finely shaded - it is spontaneous and natural in its emotional expression. The feelings of the introvert are delicate, and his tendency is not to repress emotions directly, but to delay expression or to vary it in unusual ways. Self-analysis and self-criticism are absent in the extravert, but they are present to a considerable degree in the introvert. Generally speaking the extravert is tough-minded and has a pragmatic outlook, while the introvert couples tender-mindedness with an idealistic outlook. To the extravert, the introvert appears a "sick soul", while to the latter, the extravert is often an exhibitionist and a bore.

This sort of characterization is an example of "ideal types", such as that of Spranger. Only straightforward cases are described. Jung did assert that his psychological types were basic, but generally they are not accepted as ultimate types, but as the extremes of a normally distributed variable.

Numerous tests have been devised to measure this variable, almost all being of the questionnaire type. The general conclusion has been that the variable can be scaled with/
with a certain degree of success, but that the discrimination is too coarse, to be really helpful. Often extraversion implies several traits within itself. Some refining of the concept was inevitable. J.P. and R.B. Guilford, for example, showed after factor analysis that quite independent clusters of responses were involved in tests for extraversion-introversion and in place of one variable they postulated three - Personality Factors of social extraversion, emotional dependence and masculinity-femininity.

So far we have noted the more well-known of the descriptive typologies. These have constituted the qualitative approach to personality and temperament.

Before we go on to pick out what appears to be important in the English and American psychometric study of temperament, it is useful to remark upon the fundamental difference in approach between those typologies we have discussed, and the American and English emphasis on isolation of single traits coupled with what has become known as the statistical approach.

The methods of a psychologist's research are determined to a great extent by his own personality, his training (the laboratory in which he works, for instance), and his own particular interests. Of potent influence, too, are the opinions of the school of psychology, if any such exists, where the research worker is occupied. The ways of thinking of a whole generation may reflect themselves in his outlook, as/

as for example the general belief in progress which glowed through the scientific thought at the end of the 19th and the beginning of the 20th centuries.

Vernon (1) has devoted an article to an outline of the differences between the German and American approaches, emphasising the essentially qualitative and clinical aspect of the German research, and the psychometric standpoint of the American. Even where some of the problems attacked have been the same, the methods used are so different that the results obtained are frequently in direct opposition. The Americans, convinced of the futility of achieving exact results comparable to those of physics and chemistry in the establishing of laws governing personality, hope at least to be able to establish statistical tendencies, based on large numbers of observations, the large numbers cancelling out the errors due to chance. Concentration on averages thus obscures the significance of the individual reaction. Thus only differences between performances in tests can be studied, not differences between individuals.

Part of the American outlook, also, is the notion that general traits are unreal. A subject's reactions to a number of tests are not determined by general traits but by many specific and independent responses. If any tests do show overlapping results, this is put down to the common elements in/

in the test situation.

The final reply is always in terms of correlational methods. Unfortunately it is not often remembered that what one puts into a test, one often draws out again, that is, that the test or rating scale may be based on a sound psychological hypothesis in the first place, and the correlation coefficient is merely confirmatory.

It would be absurd to deny the value of psychological and educational tests, but even yet both in England and America, there is far too indiscriminate a use of "tests", without sufficient knowledge of what the tests are actually testing.

It is interesting to note that America, so largely responsible for the testing movement, is characterised philosophically by pragmatism, and psychologically for some time by the Behaviourist outlook.

The German approach, on the other hand, while recognizing the complexity of human nature as do the Americans, tends to consider every piece of behaviour of an individual subject as related to the subject's total personality. If statistical methods are used, they are likely to be elementary, and the samples of population are apt to be small. If tests are used, they do not measure so much a trait, as a variety of psychological factors which must all be taken into account.

(1) Stern's "Personalistik", Klages' "Ausdruckslehre", and Spranger's psychology are the chief systems which refer every act/ (2)

act back to the total personality. Their alternative to the psychometric procedure is that of erecting psychological types, which are logical generalizations, rather than empirically determined.

German diagnostic methods have usually been regarded as too subjective, and too much based on intuition. This criticism has some justification, for there has sometimes been a tendency to decide on the interpretation beforehand, then to look for those traits, and of course to find them. This holds with interpretation of handwriting, for instance.

We are of course talking of German psychology up till 1935. In the most recent years psychology has lost its disinterested status, and has merely been exploited as one of the immoral means employed to further the "Machtpolitik", justified philosophically by Rosenberg on a line of thought through Fichte, von Treitschke, and the renegade Englishman, Houston Chamberlain.

The very short and inadequate account of both the American and German approaches we have given represents the extreme attitude of both. It is probably true to say that the psychologists with greatest insight on both sides, see something of value in the "other view". Spearman (1) for example, has said "The intuitionist tries to make ideas work without mathematics. The psychometrist, mathematics without ideas. When will both learn that two legs are better than either alone"?

Perhaps it is because of its position, somewhere between the German and American points of view, that the British psychology of personality and temperament (with one or two exceptions/)

(1) The Battle between 'Intuitionists' and 'Psychometrists'. B.J.P. Vol XXIV Pt. 4 April 1934.
exceptions) has not produced much outstanding work. Possibly
the position of sitting on the fence does not lend itself to
strong lines of independent research. On the other hand,
such a position ought to favour the combining of the qualitative
and quantitative approaches so that accumulated research work,
making use of exact means could attain psychological ends of
both practical and theoretical importance.

We shall now discuss the work done with this aim in view.

The Isolation of Traits and the Statistical Approach.

Psychologists in England and America have attempted to
devise more objective methods in assessing Temperament, and
especially they have tried to evolve more quantitative methods.
Those who think of Temperament as too essentially dynamic to
permit direct measurement, have sought to obviate the difficulty
by measuring single "traits" and investigating their statist-
ical relationships.

This standpoint has the virtue of being less open to
the inaccuracy of subjective speculation, and it lays very
much less emphasis on intuition than does the descriptive
approach. On the other hand, there is the problem of deter-
mining what sort of evidence ought to be accepted as proving
the existence of a trait. As we have already pointed out, the
mere postulation of a trait and the framing of certain questions
designed to measure it, does not prove its existence as a
trait/
trait. Some so-called 'traits' have so superficial a basis that they cannot be taken at their face value. The existence of a trait may be inferred only if it emerges from data obtained from carefully devised batteries of tests or rating scales; the quality investigated, also, must appear in a wide variety of objective situations.

The meaning attached to Temperament determines the 'traits' which are to be measured. As we have already noticed, in the narrowest application we would be measuring degrees of excitability; but in the widest sense it would be a matter of estimating the extent of 'introversion' or extraversion. Only indirect methods are possible in measuring such qualities apart from that of excitability.

As a result some psychologists have taken as their entities, not the traits appearing at the level of behaviour, but underlying 'factors' assumed to be common to a number of such traits. First of all ratings are obtained on a number of traits, and the existence of factors is then established by statistical means. For instance R.B. Cattell(1) defines as "surgent" and "desurgent" two groups of temperament traits which have an evident relationship to Jung's extravert and introvert types. Introversion is equivalent to "desurgency" plus a certain kind of emotional maladjustment. Surgency and desurgency are determined by the presence or absence of a 'general temperamental' or "C" factor. The related temperamental traits are:

### C +
- Surgent or Extravert
- Cheerful
- Natural
- Sociable
- Humerous
- Adaptable
- Quick of Apprehension
- Hasty
- Confident

### C -
- Desurgent or Introvert
- Solemn
- Formal
- Unsociable
- Earnest
- Conservative
- Slow of Apprehension
- Introspective
- Anxious

Cattell's "C" factor correlates highly with the factor of "fluency of association or "f" factor, postulated by Bernstein. Cattell's tests include interpretation of blots, the completion of forms, and directed verbal associations. The results must therefore be partly influenced by intelligence and special abilities and should be interpreted in this light.

An analysis of temperament has been done by Burt. He analysed the overtly displayed temperament traits of 124 neurotic and delinquent children. The first factor emerging from the analysis, being a general one, Burt postulated "general emotionality". Two further factors he designated as bipolar, one being aggressive as contrasted with inhibitive emotions, and the other pleasurable as contrasted with un-pleasurable emotions. He sums up his position as follows, in a symposium on Instinct, recapitulating the tentative conclusions which he published in an earlier article on Character and Personality.

(3) The Factorial Analysis of Emotional Traits, Character and Personality VII.
"All emotional tendencies prove to be positively correlated: we have, therefore, to recognize an all-embracing factor of general emotionality. When we have eliminated this general factor, we are still confronted with two or more broad temperamental factors, each covering a large field of behaviour: of these the most conspicuous is a factor apparently sorting nearly all emotional reactions into two main sub-divisions - an aggressive or sthenic, and an inhibitive or asthenic; a further factor is discernible, cross-classifying them into another pair of sub-divisions, namely, those chiefly attended by pleasure and by unpleasure respectively. The discovery of these more comprehensive factors, therefore, lends considerable support to those who maintain that personality forms a more or less integrated and consistent whole.

But when these 'common factors' have been ruled out, a large number of more specific correlations become apparent. By their arrangement they suggest that the numerous stimuli and the still more numerous reactions may be sorted into about a dozen or so loosely demarcated, partly overlapping clusters. These clusters in their turn correspond fairly closely with the several instincts that appear in the lists compiled by earlier writers.

Burt's conclusion regarding instincts is that they may be looked upon as innate, conative, group-factors. They are group factors because they enter into some actions only. In his view, the usefulness in postulating such abstract factors/
factors is that "the relations between the various stimuli on the one hand and the reactions, on the other, are extremely loose and variable." There is probable prediction, not mechanical certainty.

Whether or not factorial analysis will prove to be the instrument of reconciliation between the subjective and objective points of view in psychology remains to be seen.

When we speak of "traits", it would probably be best to regard them as qualities of the personality as a whole, and not as applied only to temperament. We are therefore not confined to any particular level of origin; and we are not postulating structural entities, but merely convenient means of classifying behaviour qualities whether observed in one personality only, or as common to many. G.W. Allport (1) defines as 'common traits' those "aspects of personality in respect of which most mature people within a given culture can be compared." When using such a definition it is unnecessary to discuss the question of the innateness of traits; for practical purposes they may simply be accepted as dispositions forming part of the personality.

Our sole evidence for the existence of a trait, also, is a tendency towards a certain overt quality of behaviour, but the appearance of the same trait, even to the same degree in two individuals need not be due to the operation of the same psychological mechanism in each case. One talkative subject/

subject may be giving expression to a care-free, extraverted mentality, while what is apparently the same talkativeness in another subject may be a compensation for some inner anxiety.

Complete studies of individual personalities cannot be made merely by assessing traits; for such studies the methods of the psychiatrist or clinical psychologist, which stress the examination of psychological mechanisms, and differentiate between inherent and acquired factors, are essential. The main value of the assessment of traits when studying the individual, is with reference to the group of which he is a member, and the extent to which light is thrown on the distribution of a certain behaviour quality within the group.

We see, therefore, that since an individual must be assessed relative to his group, cultural differences between the groups themselves must be taken into account, and interpretations must depend on those cultural differences. The expressive use of hands normal in France, would be considered to be indicative of excitability in this country. Similarly the average educated English woman's voice would tend to sound more than usually loud to the woman inhabitants of a remote island in the Outer Hebrides.

Current Methods of Assessment.

An assessment of Temperament cannot be expressed in a single figure in the manner of the Intelligence Quotient. No measure is as direct as that of the intelligence test.
We shall now discuss the principal methods in use within recent years.

Measurement of Traits.

Certain psychological investigators decided that the best way to attack the problem of studying Temperament experimentally was by isolating single traits. Such a trait is "persistence".

One of the main attempts to treat the temperamental problem in this way was by T.H. Howells. He assembled a battery of tests designed to measure persistence. He measured the time during which subjects submitted to various situations of steadily increasing physical discomfort. The items included persistence in holding a dynamometer, set to half the subject's maximum grip, in spite of slowly increasing pressure by a needle on the palm of the hand and by wedges pressing upon bony prominences and soft tissues. Subjects were also offered the opportunity of taking varying chances of a severe electric shock in order to obtain a supposed increase in score on the battery as a whole.

Howells' results were for the most part significant. His score correlated with ascendance in the Allport Ascendance-Submission scale $0.44 \pm 0.08$; $0.37 \pm 0.06$ between persistence and religious radicalism. The multiple $r$ between academic status and a combination of intelligence and persistence scores was $0.64$ as against $0.51$ for intelligence alone.

The/  

The difficulty arises, however, with Howells' results, as with those of many others of this nature, of the extent of the influence of motivation on the results. In Howells' research the situation was a purely competitive one. The subjects were students and they were told that the results were to be posted in order of rank. Subjects who are aggressive in social situations, and those who, while egocentric, are capable of good social adjustment, would probably tend to score very highly in these circumstances, whereas if a guarantee of secrecy had been given, quite different results might have been obtained. Persistence, as a trait, must depend to a great extent on the situation in which it is to be exercised. It seems to the present writer that a result such as Howells' would have to be confirmed by some other measure where the part played by motivation would be controlled in another direction, that is, it would be necessary to devise a test or tests to measure persistence in which the aim is not immediately apparent to the subjects. The moment the competitive factor is introduced into the situation, it follows surely, that the test situation itself undergoes a change.

Another earlier attempt to measure persistence was made by Fernald.\(^1\) He devised a Kinetic Will Test in which he made his subjects stand on tip toe facing a dial which recorded the height of the heels from the floor.

Similarly/

Similarly A.F. Bronner measured the time a subject could hold heavy dumb-bells when the arms were extended level with the shoulder. This really measures self-respect and pride in endurance, with a small element of physical ability.

Hartshorne & May made an attempt to measure persistence which seems to come nearer the mark. They set children to do exceedingly difficult puzzles, the score being determined by the length of time during which the subject persisted in his almost hopeless task.

In one way, at least, this appears to be a better measure of persistence, from the point of view that it is indirect; or to put it in another way, persistence will appear as a by-product, since the initial aim of the subject will be the "doing" of the puzzle.

One of the best-known traits to be investigated by itself has been perseveration or "p". The great difficulty here is that perseverance is directly dependent on a simple neural function, and is therefore, in accordance with our previous discussion, essentially a temperamental quality. To quote Spearman, also, - "There is no relationship between persistence, which is volitional and conative, and perseverance which is that neural quality which determines the manner in which such 'psychic' energies as may exist are discharged".

Its nature is cognitive and Spearman looks upon it as expressing the inertia of his postulated nervous energy.

Various types of activity are usually included under the term perseveration, of which the main forms are the following:

1. Persistence of sensory stimulation after the stimulus has ceased to act, which constitutes the physiological basis for the illusion of movement in the films; or for the 'fiery' circle when a glowing match is whirled rapidly.

2. The apparently spontaneous emergence of an idea after it has been some time out of mind, without a renewal of the corresponding stimulation e.g. when a tune keeps running in one's head.

3. The subconscious or purely physiological continuance of effects of experience which gives us what has been called "perseverating determining-tendencies".

Lankes, before his experimental investigation of perseveration, raised the following two questions:

(a) Is perseveration a general factor?

(b) How is perseveration, as a peculiarity of the cognitive side of mental life, related to the "perseveration" qualities of character i.e. perseverance or persistency of will?

Lankes gave an affirmative answer to (a) with the following qualification. "The lowness of the single t's tends to show that normally with average, unselected subjects, this general factor/

factor, though present and operative, is not very strong."

To (b) he gave a less final answer. Contrary to Lankes' expectations (but not to ours, in the light of more knowledge on the subject!) the perseveration quality failed to correlate or correlated to a slightly negative extent with persistence qualities of behaviour as estimated by well qualified judges. He concluded that the will is independent of perseveration so that the self "can modify and directly counteract its own nervous system and its innate tendency towards perseveration or the opposite." He did find that the highest degree of persistence was actually found in conjunction with perseveration, but that a high degree of the former could exist without the latter being extreme.

The method of testing "p" devised by Pinard (1) and modified by Cattell (2) and Stevenson (3) is one of the best known. The tests utilized vary slightly but all are based on the same principle. A subject, for example, is asked to write the capital letter S in rows for 30 seconds, then to write a reversed 2 for a further 30 seconds, then to alternate the normal and reversed forms. A similar test is with the normal Δ and a reversed triangle V.

Conflicting results have been obtained from these tests and their statistical consistency and validity is not high. It is fairly certain, however, that very high and very low "p" scores are symptomatic of personality difficulties.

High "p" is generally associated with introversion, and low "p" with extraversion.

Those rather indeterminate results have been accepted by some psychologists as a kind of "proof" of the non-existence of p. What really seems to be the trouble is, however, not so much the non-existence of such a factor, but the very inadequate means used to assess it. Obviously such tests as the reversed S and A type are very much subject to acts of will on the part of the subjects. If perseveration is an innate neural quality, some purer type of test must be utilised, or else the testing material and periods must be prolonged to a much greater extent. If perseveration is a general factor, it must be present at sensory levels as we have already noticed. Indeed it is held to influence the threshold for flicker, and attempts have been made to test p under such conditions.

The present writer is of the opinion, also, that this is one of the few instances in which the criticism against the testing of 'simple' sensory levels as artificial and psychologically meaningless to the subject, does not hold. It would be interesting, for example, to estimate individual differences in after-images in a perceptual situation such as that in dark-adaptation experiments. The present writer herself experienced fairly frequent after-images under such conditions, while another experimenter under the same conditions was aware of scarcely any at all.

In our discussion of measurement of traits we can scarcely pass over the pioneer work done in 1915 by Webb. His categories/ (1) Character and Intelligence. Mon. Suppl. B.J.P., Vol. I.
categories are mainly ethical, and although we did say that
the investigation of character did not concern us, because
of its ethical connotation, we must, however, make some mention
of Webb's research, because of its pioneer significance.
His extensive study with ratings has somewhat diminished in
value because of the very subjective aspect of the assessments,
and because his categories are not necessarily constant.
Take honesty, for example; the ethos of many a business man is
often a peculiarly inconsistent affair when judged according
to logical values. It may however, be psychologically per-
fectly consistent with his own character, and with his adapt-
ation to his own group. Along with changes within the group,
a corresponding "trait" adjustment may take place within the
man himself, so that it is rather unwise to attach much
significance to such fluctuating categories, as far as our
present purpose is concerned. Sociologically, of course,
such traits are of decided importance.

Webb, from his statistical study of "character and
intelligence" claimed to have established a general character
factor "w" which appears to correspond with the degree of
personality integration of the subject. "Its nature is best
conceived", he said, "in the light of our present evidence,
to be in some close relationship to 'persistence of motives';
that is, to depend upon the consistency of action resulting
from deliberate volition, that is from will. It thus appears
to coincide more with Ach's conception of will than with
either/
either 'perseveration' or 'secondary function'. Further evidence is necessary."

Certainly in Webb's work we are likely to have as reliable a rating of other people as was possible at the time, but the many points of inadequacy in rating scales were not then fully realised. The "halo" effect, for instance, was not taken into account, such as the influence on the judges of conspicuous physical and social traits. The well-known case of Captain X (academically distinguished, but unpopular with his colleagues) as reported by Rugg in his analysis of personnel work in the Army during the 1914-18 war, illustrates the strong teniency of individuals to be governed in their judgment by the general mental attitude towards a given individual; and this general mental attitude may be due to a "conspicuous" trait of the individual concern. The main importance then, as we have said, of Webb's research is that it was a noteworthy first attempt to measure on a large scale in this field, but the results have diminished in value because of the errors incidental to the method.

A few years after Webb, June Downey published in America a book giving an account of the tests which she had devised to assess the Will-Temperament, as she called it.

She/


(2) The Will-Temperament and Its Testing, 1924.
She discusses several motility variables and makes use of the psycho graphic method. The present writer is of the opinion that in spite of the not very high reliability of some of her tests, and the artificiality of some of the situations presented to the subject, Downey's approach was of more value than is often allowed. Many of the criticisms of her work appear also to have come from a certain misconception as to the nature of her variables.

Downey's approach to the Will-Temperament followed upon 2 previous lines of investigation.
1. An attempt to determine whether muscle-reading could be developed into a scientific method for the study of individuality. (1)
(2) An investigation of the degree to which personality could be determined from exercises in handwriting.

Downey had been interested in muscle-reading i.e. the kind of mind-reading which is really determined by unconscious movements, for it demonstrated to her the extent to which man reveals his thoughts, his decisions, his emotions through involuntary movements.

Finding that muscle-reading depended too much on the skill of the operator, Downey changed to the investigation of handwriting. In this she claimed that it was possible to identify in handwriting signs of motor explosiveness and motor inhibition. This was to be determined by the production of automatic writing and by the characteristics of handwriting/

handwriting that is deliberately disguised.

Her use of "Temperament" refers to an innate, relatively permanent disposition. "Will" is prefixed to indicate the nature of the particular disposition that is tested. This is done to make clear the fact that the test is not concerned with emotional reactions. "Will" refers to the dynamic pattern of the individual. The will-temperament test since it taps chiefly innate tendencies is only in part to be conceived as a character test.

Downey followed others (James, for instance) in stressing the level of activity or the energy output as a fundamental phase of personality. She devised certain short tests which in turn determined the extent to which the following traits operated within the individual.

1. Speed of Movement: This to be assessed relative to size and age of testee. (Test - average of 2 trials writing United States of America).

2. Freedom from Inertia: Tendency to work at one's highest speed without external pressure; little tendency to relax speed; quickness in warming up to a task. (Ratio of natural speed to capacity speed.)

3. Flexibility: Ease and success in readjustment; capacity to modify one's routine reactions. (Test - disguised handwriting and copying of model hands).

4.
4. **Speed of Decision:** Quickness in reaching a decision or conclusion. A slow reaction may be due to caution in weighing the elements involved in a situation, or it may be caused by one's being side-tracked by irrelevant matters or by a rambling procedure. (Test - checking own traits from a list.)

5. **Motor Impulsion:** This trait refers both to impetuosity and energy of reaction. It denotes the ease with which brakes or inhibitions are removed, and also the tendency to an explosive reaction when the brakes are off. (Test - speed of writing maintained under distraction of attention.)

6. **Assurance:** (Later called reaction to contradiction): This refers to the degree of confidence with which one maintains one's opinion against contradiction. (Test - To choose easy and hard mental test in 2 envelopes, both actually being blank.)

7. **Resistance:** The vigor with which one reacts immediately to a blocking of one's purpose. It ranges from a strenuous reaction, to complete passivity in the face of opposition. (Test - blocking movements of writer's pencil.)

8. **Motor Inhibition:** Capacity to keep in mind a set purpose and achieve it slowly. It involves power of motor control, imperturbability and patience. (Test - time for which writing U.S.A. could be voluntarily retarded.)
9. Care for Detail: Attention to details. Not equivalent to accuracy, which usually carries some implication of power of keen analysis. One may possess great capacity for detail and yet lack penetration in the selection of details. (Test - exactness with which a model imitated with instructions to copy exactly; accuracy with second model with only general instructions; relative times of rapid and exact imitations.)

10. Co-ordination of Impulses: Capacity to execute a double task without a preliminary trial; capacity to handle a complex situation successfully without forgetting either factor involved. This trial is probably allied to keep one's head in a confusing situation. (Test - Success in writing in narrow space at top speed.)

We have listed the traits in slight detail, because although the present writer had no such list in mind (the Downey book having been looked at by her very cursorily some two years beforehand), some of the traits, especially the pivotal ones - motor impulsion and inhibition - are similar to those which she herself observed later on with performance tests.

Undoubtedly Downey over-interpreted the results of her tests, at least as far as their general application was concerned. It is quite likely that her own personal interpretations were very sound, but that she used a larger number of factors in her estimations of subjects that she actually set/
set out for other investigators with her tests. Obviously she herself was a good appraiser of temperament, since she was keenly aware of the many stumbling blocks in the way of objective estimation, such as "halo" effects, already established intuitions waiting confirmation, and the most important factor of all, in the present writer's view, - the very great tendency of the investigator to make his estimates relative to the needs of his own temperament.

Downey scored each trait on a 0-10 basis and erected a profile for each individual tested on the basis of the traits. She found that highly patterned profiles were infrequent. The pattern of the profile was a matter of the relative emphasis of different traits in the same individual. Downey's own focus of interest in the profile centred in studying the inter-relation of traits in an individual, rather than a comparison of scores for different individuals. She made little use of the total score obtained by the sum of the scores on all tests, for some of the traits must of course be negatively correlated. On the other hand a very high total score undoubtedly would reveal a forceful personality, and a very low score, a weak or colourless or psychopathic personality.

One virtue of Downey's approach is similar to the one we found in Heymans' classification, namely its dynamic aspect, for within an individual can be many different combinations of her variables. In a general way, too, she resurrects the 4-fold classical scheme in terms of her own -

Sanguine/
Sanguine temperament (quick and weak reactions.)
Choleric temperament (quick and strong reactions.)
Melancholic temperament (slow and strong)
Phlegmatic temperament (slow and weak.)

Her opinion on 'transfer of habit' is also interesting in view of the specificity which we already observed to be characteristic of the American outlook. Her opinion is that a specific habit will have merely a specific reaction, but she qualifies this by her view that when a specific habit "plays into the native tendencies", we may get it generalized.

**Verbal Associative Process.**

Emotional factors exercise a selective function which can be demonstrated in associative processes. An associational element occurs in practically all tests of emotional qualities, including Cattell's tests for surgency and the Rorschach method. The basis of most of these tests is of course Jung's analysis of the associations as contained in the records of psychiatric patients in response to selected stimulus words. G. Kent and A.J. Rosanoff compiled a more general form of association test using a standard list of words. Although the Kent-Rosanoff list, being thus selected, lends itself to more rigid experimental conditions, the assumption is made at the same time that the different significances of the words are universal. One would need to have an enormous amount of data from various cultural groups to shed more light on the reality of this universality. Apart from this disadvantage, however, (though there/
there certainly seem good foundations for the interpretation of "complexes" of the most obvious sort) the word-association method must be one of the most valuable of the simply administered clinical methods. It also has great group-testing possibilities, as well as being very fruitful in individual testing.

Intelligence Test Scores and Behaviour during Tests.

Clinical and educational psychologists must often have thought how much qualitative information of the temperamental sort could be obtained on the basis of intelligence test results. The individual scatter within the test, or battery of tests, coupled with individual behaviour reactions during the test, must certainly have some significance worthy of careful investigation.

There has been very little written on this subject, but it is probably fairly safe to assume that many psychologists have not relied on the bare intelligence score, but have estimated other qualities manifested by the subject, during the same intelligence test. Emotional qualities, for example, are extremely difficult to assess directly, and may possibly be indirectly estimated in the general testing situation.

Attributing emotional significance to the distribution of test scores, also, assumes a dynamic point of view, which is often neglected in the orthodox psychometric approach. Such scatter may therefore be an indication of stability of personality,
personality, or of certain personality types.

F. Mateer (1) considered the scattering of successes and failures through 4 or more 'mental years' in the Binet scale to be indicative of psychopathy and other writers have expressed similar views.

Z. Piotrowski has claimed, however that it is the pattern rather than the degree of scattering on the Binet scale which is characteristic of emotional disturbance. A more promising approach is probably one based on the results of a battery of specific tests, rather than of the items of a mixed scale such as the Binet. The representation of the results would be in profile form. It is likely that the more unstable types might fail on certain tests, not because of lack of intelligence or special ability, but because of their difficulty in concentrating, their inability to perish, their lack of capacity to inhibit emotional responses irrelevant to the situation.

C. J. C. Earl (2) in a study of adult morons found that when there was a marked preponderance of verbal ability over performance test ability, the individual tends to be unstable with a neurotic preference for symbols rather than for realities. When verbal scores were markedly lower than performance, a 'verbal neurosis' was believed to reflect early educational difficulties. Earl of course does not take special abilities into account here, although in the main, his thesis is probably sound.

The/
The disadvantage with direct interpretation of test behaviour is of course its non-quantitative aspect, since it is based mostly on fleeting impressions and the rapid interpretation of momentary phenomena.

**Analysis of Psycho-Motor Factors.**

Psycho-motor behaviour is a relatively direct and simple expression of neural function, and through muscular tension and a variety of movements it provides clues to personality. Descriptive phrases and words such as "shifty eyes", and "wriggling with excitement" show to what extent we make use of such clues in our daily life when we assess our fellows.

Attempts to systematise those movements have been made with varying degrees of success. Allport and Vernon used a large number of tests with 25 young adults and noted the movements involved in reading, walking, tapping, stylus drawing, and the estimation of size, weight and distance were also analysed. Dynamometric measurements of grip were taken, and writing and resting muscular tension studied. On the basis of careful statistical examination of their results, Vernon and Allport concluded, in line with psychologists who have made a descriptive approach to personality, that "no general factor which directly and unambiguously underlies each or every act" could be found. Instead there seemed to be general tendencies and dispositions. "There are degrees of unity in movement/

movement just as there are degrees of unity in mental life and personality."

Their results confirmed those obtained by other workers using tests of more limited scope and involving single psychomotor phenomena. The resemblance in methods of measuring muscular tension whether during muscular activity or in a testing state is the closest.

These results of Vernon and Allport's are of significance because of their application to normal subjects. So much material of this sort is obtained psychiatrically, and the relevance to normal subjects has often been questioned. Also the tests were extremely carefully administered and the results treated in a more exact way than is usual with this kind of data.

Working with psycho-neurotics and emotionally maladjusted subjects, E. Jacobson (1) showed that there was a relationship between a general increase of muscular tension and emotional difficulties of various sorts. He claimed, even, that emotional difficulties could be overcome by education in relaxation.

Muscular tension was the basis for an experiment by E. Duffy (2) with young children. She placed a bulb dynamometer in the subject's idle hand during the performance of tasks of reaction-choice type. Her results showed that with both high/

---

(1) Progressive Relaxation. Chicago Univ. Press. 1929. (Quot. from Earl)
(2) "The Measurement of Muscular tension as a technique for the study of emotional tendencies". Amer. J. Psych., 1932, XLIV.
high and irregular muscular tension, personality difficulties were to be found also. The same subjects also tended to perform the experimental task less well, which Duffy put down to the increased muscular tension. This conclusion does not seem to be warranted, however. Increase in tension was more likely to have represented the spread of excitation in the nervous system as part of a reaction to a situation of difficulty. This at any rate, is Earl's view of the increase in muscular tension under such conditions. In cases where the rise in tension was irregular and unduly high, the lowered efficiency in the performance of the task would represent the psychological aspect of the "over-reaction" of which the muscular tension is the physiological expression.

An interesting theory has been put forward by F.I. Golla and S. Antonovitch in a similar piece of work on muscular tension. The increase in muscular tension with the performance of an intellectual task actually precedes the commencement of the intellectual activity. This observation has led them to the conclusion that the increase is an atavistic remnant, a phylogenetic relic of a reaction preparatory to physical action.

This theory coincides to some extent with an observation of the present writer of muscular contraction of the thumb in several subjects under performance test conditions, before beginning the test.

The McDougall dotting apparatus has been used by M. Culpin and M. Smith to differentiate between certain personality/

(1) "The relationship of muscular tonus and the patellar reflex to mental work". J. Ment. Sci., 1929, LXXV, 234-46.
personality types, and it is claimed that significant results have also been achieved with the very simple tapping test devised by Whipple and Healy.

This tallies also with another observation of the present writer, while watching girls in a sack-making factory doing some performance tests. The peg-board, used as a test of manual dexterity, seemed to her to have possibilities as a simple diagnostic test of temperament. The 'reality' of different 'types' of temperamental reaction seemed every bit as great as of the manual dexterity. One girl, for example, achieved a high score relative to her group, by an "attack" of great speed, with quick exaggerated movements of the manic type. Although she gained a high score, the quality of her performance was not proportionally high, for in her exaggerated hurry she dropped several of the pegs, so that from the observer's point of view, the spectacle of one or two loose pegs rolling about the board while the others were being put in at an almost frenzied rate, did not give the appearance of manual competence which the bare score achieved by her would convey to some one who had not witnessed the actual performance.

The combination of psycho-motor factors with associative phenomena has been made in a method devised by A.R. Luria.(1) It is based on the theories of the modern Russian school, and having their origin in Pavlov's work, they are characterized by/

by insistence in the physiological aspects of behaviour in general and of affective process in particular. Motor phenomena, both voluntary and involuntary indicate emotional disturbance better than pulse, respiration or blood-pressure, according to this view.

In Luria's experiment, the subject keeps both hands on pneumatic receptors while a series of stimulus words is read out to him by the operator. As he gives a verbal response to each word, he depresses the right-hand receptor, but he has been instructed that his left hand should remain motionless throughout. Graphic records are made of the latent period of response, the voluntary movements of the right hand, and the involuntary ones of the left. There is also a time record, and the actual stimulus and response words are noted.

By the use of verbal stimuli of varying degrees of intellectual difficulty and emotional significance for the subject, Luria was able to produce varying degrees of disorganization in the responses of different types of individuals. The delays in giving the verbal response, and the fluctuations in the involuntary tension maintained by the one hand and in the timing and force of the voluntary pressure exerted by the other, were found to vary with different subjects and with different stimuli. The 'reactive labile' subjects exhibited these phenomena more often and to a greater degree than the 'reactive/
'reactive stable', while various psychopathic types exhibited them to a high degree in practically all the experimental situations. Neurotic and unstable subjects tended to show severe disorganization in every response, even when no stimulus was given.

Luria's method has been extensively used in Russia, but it is not in general use in other countries. Its validity in revealing the existence of unconscious emotional factors has been experimentally demonstrated by P.E. Huston, D. Shakow and M.H. Erickson (1) in a study of hypnotically induced complexes by means of a technique which Luria had himself employed.

The Rorschach Method.

Much has been claimed for the Rorschach method of assessing personality. It consists in analysing the verbal interpretations by subjects of a series of 10 ink-blots, 5 of which introduce an element of colour. The interpretations are numbered and classified into several categories, according to various factors:

1. The stimulus selected: whether it be the whole blot, some detail, or peculiar details etc.

2. The determinant and quality of the response: whether dependent mainly on form, movement, colour or chiaroscuro; also whether the form content is 'good' or otherwise.

3./

3. The content of the response: the nature of what is seen, whether human figures or animals etc.

In addition to classifying the individual responses in this way, the following general features are also scored -

I. The relationship between responses to the blot as a whole, and those to various types of detail.

II. The order in which the various types of responses to each blot occur.

III. The relationship between responses involving animal figures, and of responses whose form content is 'good'.

The Rorschach method is in direct line with the psychology of personality which treats psychical activity as an inter-dependent whole. It therefore follows that it is in direct opposition to any theory of unitary traits.

It has been lavishly praised for the great insight it gives into personality, but it has also been criticised because of the subjectivity of the estimate, for the result depends to a great extent on the skill of the administrator. Its virtue of being not confined to specific reactions, is also one of its drawbacks, for it leaves it open to widely different interpretations. It has been used over an extensive range of subjects, more especially with abnormal subjects. There is little doubt that it can be a valuable diagnostic instrument, so long as the interpretations of the results are made with restraint, and with reference to other data obtained from the individual.
Measuring Single 'levels' of Personality.

Single "levels" have in the main been measured by tests of the questionnaire type. This kind of test is so well-known that it is unnecessary to go over what is already very thoroughly trodden ground. Some reference to questionnaires will be made, in any case, in the next chapter.

Conclusion.

From our survey of the most outstanding of the various approaches to Temperament, including some of its aspects which fall within the wider field of Personality, made within the last 25 years, as well as the current methods of assessment, there are certain noteworthy points which emerge.

None of the tests or methods which have been described can be considered entirely satisfactory, although without exception, they are useful to the investigator, some of course to a greater degree than others.

What is of importance first of all is what we want to know or investigate. We then adopt the means at our disposal to the required end. If we merely wish to make a rough description of the individual the typologies are adequate, preferably one such as that of Heymans because of the dynamic aspect of the 3 basic variables.

We may, however, hold the viewpoint that the different levels of personality are so inter-dependent, and that the nature of the initial drives to action has become so blurred, that/
that we give up hope of being able to isolate any causative factors in the temperamental make-up of the individual and resort to the use of questionnaires dealing with sentiments, attitudes, interests etc. in the belief that only at the level of opinion can anything "real" and meaningful be known about the individual under investigation.

There are always some people, however, who cling to the hope that in spite of the indisputable veiling and distortion of human motives by the inter-action of cultural forces, some sort of fundamental psychological (or better, "biological", being more inclusive) drives may be revealed. We know that Temperament is never static. It may be modified or even radically changed by influences such as maturation and decline, metabolic changes, disease and psychogenic forces.

For Whichever of the methods we use, if it is to be one only, we must not make exaggerated claims for its diagnostic powers such as have been made by enthusiastic supporters of a single-method such as the word-association tests, enocrine analysis, the Rorschach test, or psycho-analysis. For some problems one method is the most useful, for others a different method or methods.

Also it is of importance that our procedure for obtaining data must be measured against some other criterion of interpretation.

And finally, what is probably of greatest importance is that we maintain a balance between rigid psychometric standards, having no psychological meaning in the field of Temperament, and/
and loose subjective assessments leading to extravagant and far-fetched claims based mainly on intuition and having no sort of adequate proof.
A. Questionnaires.

With no very clear ideas in mind as to what experimental means she was to employ in investigating the problem of Temperament, the writer decided that first of all, it would be useful to select a small group of undergraduates who might be expected to yield interesting introspective information under whatever experimental conditions could be later devised. She decided to use the Allport-Vernon "Study of Values" questionnaire as an instrument of selection of the 1st Ordinary Class in Psychology 1939-40 at Edinburgh University. This test had already been used by her\(^{(1)}\) and it was found to sort out to a fairly reliable extent the relative prominence or weakness in the individual, of six general interests or values - theoretical (scientific), economic (utilitarian), aesthetic, social (humanitarian) (the least reliable of the values), political (power-seeking in sense of Machtmensch), and religious interests. These values are based on the 6-fold classification \(^{(2)}\) of Spranger - the assumption being that these are generalized interests within the total personality. While not subscribing to the implied objectivity of these values in Spranger's "verstehende" philosophical-pychology, and thinking of them from/

\(^{(1)}\) In education thesis for B.Ed. degree.

from a subjective point of view as merely psychological characters attributed to things in virtue of somebody's interest in them, with no reality apart from such interest, the present investigator considered that they would provide a reasonable basis in a rough sense for choosing the type of subjects she wanted. She had found that adults interested in psychology and with good introspective ability, came out high in the theoretical and aesthetic values, sometimes in "theoretical" alone, but not on "aesthetic" unless combined with "theoretical".

It was very disappointing to find from an analysis of the results of 65 subjects that not one single member of the group showed a significantly high score on both the theoretical and aesthetic values. Neither did any subject show a significantly high score on the theoretical value alone, although several came out significantly high on the aesthetic value by itself.

On second thoughts this was perhaps to be expected, for the average age of the group was lower than that of the postgraduate group including as it did, members of the University staff, which had been previously tested. It was the writer's own experience that no particularly consistent "Weltanschauung" was common amongst the mass of undergraduates, but there had been the odd exception here and there, and it was hoped that a few such exceptions might appear within the Psychology class group. Unfortunately, also, the period of the test was the first year of war and the small sample of male undergraduates
(18) was mainly confined to divinity students, who of course came out very significantly high (though not all of them!) on the religious value.

The idea of selecting subjects from this source was therefore given up, but it is of some slight interest to record the average "profiles" of men and women students on the 'Values' questionnaire, since it leads us to certain conclusions about the use of questionnaires.

The maximum score on each value is 60, and a score of 40 and over on any value is very significantly high. The average score on each value is 30.

<table>
<thead>
<tr>
<th>Mean Scores</th>
<th>Theoretical</th>
<th>Economic</th>
<th>Aesthetic</th>
<th>Social</th>
<th>Political</th>
<th>Religious</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Group</td>
<td>24.6</td>
<td>21.5</td>
<td>31.3</td>
<td>32.0</td>
<td>29.5</td>
<td>40.2</td>
</tr>
<tr>
<td>N = 65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male Students</td>
<td>23.7</td>
<td>21.1</td>
<td>25.9</td>
<td>34.8</td>
<td>30.2</td>
<td>43.8</td>
</tr>
<tr>
<td>N = 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female Students</td>
<td>24.9</td>
<td>21.7</td>
<td>33.4</td>
<td>30.9</td>
<td>29.2</td>
<td>38.9</td>
</tr>
<tr>
<td>N = 47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Group Profiles on Allport-Vernon "Study of Values"
W.E.A. Group Profile on Allport-Vernon "Study of Values"

N = 10

Above 40
Very significantly high

31-28
Average

Below 20
Very significantly low

As can be seen from Graph I the average score on each value lies between 28 and 31, so that theoretically the average profile for the graph should be rather like a straight line. This is what appeared with the large American group upon which the questionnaire was standardised, but with the Scottish postgraduate group, there was a slight dip in the economic value (possibly indicative of a less pragmatic outlook!), and a slight rise on the aesthetic value. With the undergraduate group, here presented, the same kind of profile is obtained except that it is much more patterned. There is a significantly strong religious bias, even apart from the divinity students. The present writer thinks that this may quite reasonably be a "real" enough result, for the average undergraduate outlook is fairly conventional, and (at least during the pre-war period) the only belief of any consistency was a religious belief.

Compare this now with Graph II which is the average profile for 12 members of a W.E.A. Group. Here we have the type of graph to be expected from this questionnaire. Within this group, were several well-patterned individual profiles, but the average profile is more in the nature of a straight line.

To this same W.E.A. group, the Bernreuter "Personality Inventory" was given. This well-known questionnaire consists of 125 questions which simultaneously measure 4 common traits of personality: dominance, self-sufficiency, introversion and neuroticism. Flanagan (1) added two more: self-confidence and sociability.

This/

This time, however, the W.E.A. group showed great inconsistency in their replies, except with the "self-sufficiency" trait. We are not looking for perfect consistency, however, for there are many reasons why we should not find it. One personality will often be found to have contradictory traits within itself. But we do expect a certain degree of consistency: except for one member of the group, the degree of inconsistency was very high. We are, of course, talking of logical consistency. Psychologically the inner consistency of the traits may have been greater than was apparent on the surface.

We mention the Bernreuter results, however, merely to draw attention to some observations we wish to note, on the questionnaire type of test.

Adults of some education such as university students and W.E.A. groups are good subjects for the questionnaire type of test, especially the Allport-Vernon type of test. Vernon(1) points out that uneducated people tend to be suspicious of the object of the enquiry and that this probably also applies to children.

We found from the results, however, that even with educated subjects there is a big difference between the Vernon & Allport test, which estimates "attitudes", and the Bernreuter self-rating personality tests which asks for more intimate information.

† It was not considered worth while to measure the inconsistency statistically, or to apply factor analysis to the results. The writer has doubts about the Bernreuter method of scoring.

The former is "respectable" (few people mind giving their "opinions" - often they enjoy doing so), but the other test may possibly make the rater self-conscious and unwilling to reveal information that must inevitably be tinged with emotion. As we saw from our own results, the Allport-Vernon test appears to have given reasonably reliable results, but the Bernreuter Inventory yielded mixed and inconsistent estimates.

If we are interested in group-results only, some errors cancel each other out, but an influence which may affect the majority such as many of the emotionally toned questions in the Bernreuter, will lead to some distortion of the results. Even when the subjects are very co-operative distorting factors, such as rationalization, will occur. Also, as Vernon has pointed out, sophisticated subjects who are highly aware of their own emotional lives are apt to get high scores for emotionality or introversion without necessarily being more emotional than others who believe that they are free from such symptoms. Even when such personality test results show consistency, when the split-half technique has been applied, it may only mean that the subject's self-estimates are consistent, not that such estimates correctly represent the traits. A final drawback of such personality tests is the fact that they fail to reveal differences between psychotic or neurotic patients in mental hospitals and normal persons.

Our/

* This may have been due in part to the fact that the administrator was younger than any of the members of the W.E.A. group, and it was known that she would score the results. This may have led to some lack of confidence on the part of the subjects.
Our final conclusions about the questionnaire as a diagnostic instrument are that such tests are of value, only within definite limits. They are of low validity when applied to children and 'unsophisticated' adults mainly because of wide individual variation in interpretation of the questions. Their usefulness is limited to educated adults, and even with these subjects such tests as the Bernreuter Inventory are subject to such distorted results that their value is questionable. The present writer is strongly of the opinion that the only way in which they may prove useful is under individual testing conditions, where the administrator having first inspired confidence, may be referred to for help in interpretation of the questions, should such a need arise.

Attitude tests such as the Allport-Vernon "Study of Values" and those of Thurstone have probably wider applicability, though it is again only with educated subjects that any approximately "true" results will follow. Certain sociologically-important results may however be obtained from using such tests with wider groups. Group sentiments and attitudes are interesting and important sources of information.

The present investigator thinks that personality tests of the questionnaire type have now reached a stage where the idea should be given up that they are measuring real "traits", until a wide investigation of the meaning attached to the questions by the subjects and in what terms they are interpreted, should be carried out. A comprehensive list of questions should be set/
set in which a large number of subjects are asked to state what they think the questions mean, not to what degree or how far they are characteristic of themselves.

In other words, we are advocating more investigation into what the tests are testing. We shall return to this question of item analysis later, when we discuss batteries of performance tests.

B. Individual Case.

About the same time as the writer was trying out questionnaires as the initial and most obvious method of investigating Temperament, she tested a little girl of 7 years of age who was brought in from a country town to be given an intelligence test. She was reported to be extremely backward in school in every subject except arithmetic, and to be unable to speak intelligibly. It was suspected that she was mentally defective, and some people had begun to refer to her as "the loony" behind her back.

Assuming that she might possibly be 2 years retarded in intelligence development, she was given the Merrill-Palmer scale of performance tests. Apart from the Action Agent part which involves vocabulary, and which she failed completely to do, she completed all the tests in record time. Then she was given the Drever-Collins scale of performance tests on which, with the exception of the Healy Picture Completion I, she tested admirably. Even with a low score on the Healy she gained an I.Q.*/

* The question of what the Healy is actually testing will be discussed in a later section.
I.Q. assessment of 140.

The marked discrepancy between her very superior performance level and her practically non-existent vocabulary, along with the fact that she did not appear to understand verbal instructions, led to her being given a test for deafness. She was discovered to be suffering from high-frequency deafness, the most probable cause being an attack of meningitis at 18 months. The expert who tested her hearing was of the opinion, without knowledge of the performance test results, that the child must be of high intelligence, for with such a handicap he could not otherwise understand how she had made an adaptation to the world such that nobody had ever realised her deafness.*

What specially attracted the present investigator's attention, however, was the qualitative aspect of the little girl's performance. Admittedly her own little universe would be mainly one of concrete objects, and scarcely, if at all, a verbal one; this would give her greater familiarity with material similar to that of performance tests than that of other children of her age. But it could not be the only explanation of the superior "attack". It was surprising that, being in danger of being sent to an M.D. school, as well as the fact that people in her neighbourhood were beginning to think of her as abnormal, she did not show signs of temperamental maladjustment. In fact the contrary was the case. Her actual "performance" on the tests could only be described as /

* This little girl has had speech lessons from a competent specialist teacher and is now making excellent progress at school, where care is taken that she may always see the teacher's face, in order to lip-read.
as 'beautiful'. Her movements were smooth and regular, neither did they exceed the number necessary for exact completion of the tests. There was no fumbling and no anxiety. There was great care and exactness without any over-emphasis, and there was very obvious and speedy comprehension of what was required in each test. The child's handling of the material was most competent in every respect.

On the basis of conversation with the mother and a communication with a teacher who was giving the little girl special lessons, it was made very clear that the home background was a specially good and affectionate one. It then suddenly dawned on the writer that the child had no reason to be maladjusted, for she had never heard any of the things people were beginning to say about her. The parents were kind and thoughtful, hoping that what they imagined was a speech defect would come right with special teaching, and doing their best to betray none of the slight anxiety they were beginning to feel. The child's main companion was her baby sister, and since she had not been very robust since her meningitis attack, it was accepted that she could not join to any great extent in many of the other children's games.

Two ideas occurred to the writer after this testing experience -

1. Large discrepancies between I.Q.s and Performance Test scores should be highly significant. What had happened in the above case might work the other way round. A high I.Q. on/
on a verbal test, with a low Performance I.Q. would possibly indicate bad adjustment to the outer world i.e. a world of concrete objects.

2. Performance Tests generally might possibly be made the basis for some scheme of Temperament testing which would have more reality than the pencil and paper test situation.

It must be pointed out that up till this time, the writer had been quite unaware that anybody had used performance tests for Temperament testing alone, although she did realise that clinical workers must have made incidental observations, such as Stutsman advocates, while using the tests primarily for intelligence testing purposes. Later she found that the idea was not so original! Stephenson (1) had explored the possibilities of such a use for Performance Tests, and C. J. C. Earl (2) had utilized them with adult morons, relating certain types of performance to certain emotional factors.

The writer had a second reason for her choice of Performance Tests. That was the fact that Temperament traits seemed to her practically impossible to investigate in artificial isolation, that what we in everyday observation describe as temperamental factors in an individual's make-up, must be considered in terms of their inter-action with intelligence and special abilities.

C. Tests with Adult Subjects.

C. Tests with Adult subjects.

It was decided, therefore, to use a short battery of performance tests with a small number of sophisticated adult subjects (4) with the idea in mind that by observation of the actual performance, and by observation alone, it might be possible to determine some kind of standard of good quality "reactions", with which other performances might be compared. The writer wished to see if there were anything common to those performances, and to what extent they exhibited individual differences. She deliberately decided temporarily to give up reading publications in journals etc. on what had already been done in the field of Temperament, so that as far as possible she could decide on purely observational grounds what to look for in her later use of such tests with younger subjects. This must be one of the very few psychological situations in which it is legitimate to know as little as possible about other research done on one's own problem!

The tests used fall under two headings:

(1) 2 puzzles with catch or chance solutions;
    2 puzzles with possible solutions.

(2) 2 performance tests.

Puzzles.

Since we were to deal mainly with the conative aspect of mind, we thought it would be interesting to investigate in a small qualitative way creative frustration. In its proper setting, of course, its ambit is as wide as life itself. Many psychologists have studied it at various periods, from different angles/
angles. Jennings investigated the frustrated behaviour of amoebae, while Fabre and the Peckhams studied insects from this point of view. Köhler and Yerkes similarly observed frustrated behaviour in apes. It was noticed that when frustrated in an instinctive process, insect, ape and child developed instinctive adaptive behaviour. The reactions of a child vary from a simple cry to other and more subtle methods. The frustrated instinctive impulses of adults in everyday life are also worthy of consideration.* It is very likely that throughout the animal kingdom, efforts to deal with frustration have been a potent factor in evolutionary process.

(a) **Miniature Safe.**

This was a metal model of a safe with 3 small knobs and S. was told that in order to open it up, he had to find the correct combination on the dial. This meant that out of a vast number of twists and turns, one chance combination might lead him to the solution. (This was probably especially hard on two of the four subjects, for they were well aware of the facts of probability!) In no case was this puzzle solved. Each subject kept at it for 30 minutes. Each S. was asked to introspect over the whole period with special reference to obstructions. Introspection on the cognitive processes involved in frustration along with their volitional and affective accompaniment, was asked for.

Extract/

Extract from introspection of S. I.

To begin with, I'm trying to see what all these things (i.e. the knobs) do. (S. turns knob rapidly, then combines it with a corner knob).

I have a feeling that I'm being put on trial and shall feel annoyed if I can't open it..... My pride is at stake. (Confines himself to 3 knobs.)

(Pause for reflection) Going to try something I haven't done. For the first few minutes I've just fiddled about aimlessly ...... Don't think I should make a good burglar ..... It's awful being watched like this! Feel like one of the monkeys in the Zoo when it's given a piece of mirror.

(Pause. Sits with chin on hand to reflect. Goes back to 3 knobs.)

At first I was interested. Now definitely annoyed. Inferiority complex. Have got to stage now when I can't think of anything more to do .... Getting very annoyed, just turning things round and round .... Feel the solution is quite simple and that it ought to be within my power to solve it much sooner than I'm doing .... If it were my own I should go on and on at it. In everyday life I should .... I feel now that I probably shan't succeed and that I shall have to give up and it annoys me. I haven't really lost interest. I should dearly like to do the thing as much as ever.

Extract from introspection of S. II.

S. spends nearly 20 minutes trying all three knobs and concentrating well. "Feeling of frustration beginning now." (Tries to combine dial with knobs.) "Feeling of helplessness in/
in front of the enormous number of permutations and combinations of the 4 gadgets." (S. begins recital of possible combinations and stops with sigh.)

(b) Small wooden box.

This small box was simply to be opened up. It appeared to have 2 sliding lids but neither really functioned.

S. I.

(After 6 minutes of concentrated effort) I can't think what to do .... I decline just to continue waggling this about. It's too silly ... I feel I ought to be able to do it .... I feel very stupid .... This is deflating my 'ego' ....

My difficulty is that I don't know what to do next .... I don't feel its going to be successful .... I'm sure this lid can't be lifted off. It must be slid off. (This was actually the correct thing to do, but S. could not put it into operation.)

S. III.

"I can't see myself getting any further with the thing .... If I could only see a little improvement. Still I may as well carry on .... But I think I've done every thing possible. (This S. hardly looked up at all. Showed great persistence).

(c) Disc puzzle.

Small, medium and large discs were to be shifted from 2 poles to a 3rd one without ever having a smaller disc under a larger one. This puzzle was actually possible in a definite number/
number of moves, and was not of the impossible or chance solution type, as were the first two.

S. II.

(Regards puzzle for at least 'a minute)

(Sighs after some time) Have feeling of indecision whether to go to A or C .... Feeling of hopefulness as I think I'm on right track .... Pleasure .... (More sighs)

I have an impulse to lift this whole lot when you're not looking, to save time .... Slight feeling of dismay at not working out so easily as I thought 3 minutes ago .... Feeling of frustration and bafflement beginning now .... growing stronger. Never felt before that introspecting could be so interrupting. Feel that I lost the thread after I told you I was on the right track. I have the feeling that if I hadn't reported it, I might have succeeded.

(S. said that knowledge of existence of a 3/4 hour time limit had contra-suggestive effect. Thought that it hindered his activity. "Knowing that there's a time limit produces something akin to bravado. I don't think I could concentrate more but I'd probably do the puzzle quicker."

(d) **Constructive puzzle.**

Bits of wood to be put together. Time limit 30 minutes

S. I.

Don't feel particularly bright today; but willing ....

This/
This is rather infuriating. Rather annoying - I'm sure I did one of those in my youth and quickly too. I must be getting old. Ah! (S. thought he saw solution coming). Ah! (again thought he saw solution. Looks annoyed and rises slightly from chair.)... I can see what to do and I can't do it. (S. concentrates strongly and nods head as he thinks.)

.... I'm very annoyed. I'm frustrated. And I'm suffering from injured pride....

S. II.

S. spends long time regarding the puzzle.

After 6 minutes - The first feeling of frustration is beginning now. (S. begins to frown) .... Feel I may be on the right track this time ..... Feeling of frustration is becoming very strong now, almost amounting to despair, but not quite .... Sudden irrational desire to fling the whole thing away.

S. IV.

Don't like giving a thing up, once I start .... Think I see what to do. No, perhaps I don't .... I usually like this sort of thing, but I don't seem to be getting anywhere here ..... Slightly anxious ..... Feel annoyed when I can't get it out, but then I can't give it up ..... Must go on ..... 

Conclusions.

From the short extracts given above, we can see that activity, from the initial goal to the attaining of that goal (if it is attained) is a "whole". Any frustration does not disperse the tension/
tension created, but increases it. It would appear that while the tension during the working of a task is expressed conatively, where there is frustration the tension often continues and is expressed in affective conditions.

It was also interesting, in the light of the research of the Würzburg school to sort out the different periods of the testing situation.

(1) The Fore-period during which S. was aware of the instructions, was thinking about the coming problem, was generally speaking emotionally "toned-up", alert and willing to put forth effort. One assumes also that some muscular sensations, not reported by the subjects, may have been present in this fore-period.

(2) Perception of the puzzle stimulus.

(3) The actual working out of the puzzle.

(4) Obstructions.

(5) Stop signal.

(6) After-period.

The resolve to concentrate was important for the fore-period. This puts us in mind of Ach's "determining tendencies", for the initial resolution of all the subjects was the determining tendency for the experiment as a whole. It was interesting to notice what happened when later obstacles clashed with determining tendencies, and also what occurred when two determining tendencies themselves were in conflict e.g. a determining tendency set up by the resolution to try to solve the puzzle, and another resolution to stop whenever the time limit was/
was declared to have been reached.

We see also from the introspection that there are cognitive, conative and affective aspects of each stage of the puzzle-solving.

Introspection under such conditions is interesting for any light it may thrown on complexes. The affective accompaniments of frustration are relatively permanent. The affective state does not pass away leaving no mark, but tends to persist and to be reproduced on the presentation of stimuli similar to those creating the original frustration. This is of course balanced by the effects of success. Even the temporary feelings of success with the puzzles in our subjects show this and this was much more manifest with the performance tests, proper, which our subjects did very competently. By the balancing and blending of the two the emotional life of the individual is obviously in great part created. Present success or frustration is a factor, but the affective condition in the present state is largely due to the experience of the past which colours all present success and failure.

The obstructions, also, had 3 aspects.

Cognitive obstructions or frustrations.

Failure to see relations between parts of the puzzles; awareness of the time factor; lack of insight.

Volitional and Conational obstructions.

Orectic obstruction because of the difficulty of the puzzles; appearance of a new volition; sense of futile striving, and lack of direction in effort.

Affective/
Affective aspect.

This was the most striking aspect. Emotionality itself a frustration, negative self-feeling, feelings of despair and hopelessness, disappointment, and annoyance* were all present to some degree in the introspections.

Muscular sensations.

These were not actually reported, but the experimenter herself saw occasional signs of muscular tension.

Performance Tests (Ferguson form-boards and Passalong).

These two tests were more satisfying to perform for the subjects, for this time they knew that a solution was possible and marked success effects were apparent. The two main results of observation of subjects here come under two headings, namely, (a) types of movement (b) method of attack.

Types of Movement.

In all subjects there was marked control of movement i.e. there were no unnecessary moves beyond the bare necessity of the tests. There was no other way exhibited of placing the pieces of the formboard except that of putting them down exactly and with precision. Some of the subjects were rather slower, or quicker as the case might be, than others, but neither type was over-marked. There was no impulsiveness, and no outstanding evidence of inhibition. There were no fluctuations in work-type. Certainly/

* These primitive reactions such as anger and rage were not in evidence. No doubt they might occur with less sophisticated subjects. In only one instance did a slight show of anger indicate a pronounced frustration.
Certainly the form boards were not difficult enough to disorganize well integrated adults, the only obstruction being the last 2 items of the Passalong test, but the writer was mainly interested in what seemed essential in the general performance; that appeared to be something related to types of movement. This, she decided, would be one of the pivots of subsequent observation with performance tests.

Method of Attack.

The method of all subjects was characterised by planfulness, and by paying intelligent attention to obvious clues. To illustrate the latter point, we noted that the pieces first chosen in the Ferguson form board were the most obvious ones with rounded edge. Trial-and-error was conspicuous by its absence. The general impression was of purposive and deliberate attack. There was conspicuous evidence of "balance".

In part, maintenance of balance is an outcome of adjustment to one's environment, and adjustment is obviously more difficult in some environments than in others. One of the really vital questions about any individual is how much strain he can stand without losing balance. In modern warfare this is of especial importance.

In conclusion, one point of fundamental concern to us, is to what extent we may consider the attitude of the subject to the tests in a narrow sense as a microcosm of the attitude to life. There are undoubtedly some situations in life which may have acquired a special emotional charge for some reason and may/
may therefore produce entirely idiosyncratic reactions. We can only hope, however, that the performance test situation will prove to be sufficiently impersonal. The problem consists essentially of success or failure and it is reasonable to suppose that affective-instinctive responses should be easily inhibited, and that in consequence the emotional reaction to the exercise of intelligence during performance tests should not be specific but should indicate a general tendency or tendencies.
CHAPTER III. CONSTRUCTION OF NEW PERFORMANCE TEST. (P.C.C.).

The writer had two reasons for wishing to construct a new performance test. One was that the drawback of most of such tests which she had come across was that they did not last long enough, at least not long enough to show persistence of effect on the part of the subject. She therefore decided that this would be the first essential. The second reason was that she wished to have a test which would estimate, if possible, the "goodness" of an individual's 3-dimensional perception. The investigation had a theory that space perception might be influenced by temperamental factors. She had watched young adults doing some tests in a factory, and in some cases there appeared to be actual difficulty in seeing the cubes in the cube-counting test. (1) One girl said that she just could not count the cubes and apparently was quite unable, also, even to distinguish one cube from another. In fact what appeared to be the difficulty was her inability to see the cubes as other than a flat figure with lines running vertically and horizontally. This particular individual was highly excitable, and the present writer wondered at the time whether, because of a certain habitual excitability, she was not preventing herself from seeing adequately in three dimensions. It looked almost as if the temperamental ill-balance were having a "pulling-in" effect. On the other hand, the simple explanation that subject's poor spatial discrimination might cause her to become excited as a result of lack of success with the test, was possibly all that was needed to explain her zero score.

What?

(1) As used as an item in the American Army Beta Scale, etc.
What particularly attracted the writer's attention was the fact that this particular test was in use as an intelligence test, and yet it seemed to have quite a big special factor loading. It would be interesting, therefore, if a performance test could be devised which would combine both the lengthy time factors and something which would involve what has become known as k-factor.

After investigating various puzzles and similar material for some time, the writer was very kindly lent, what was considered to be a puzzle of considerable difficulty, by Dr Mary Collins. It consisted of six bits of notched wood, which when put together, would form a 3-dimensional cross. The writer decided that this puzzle could probably be made into a graded test which would fulfil both of her conditions.

The test would involve a purely perceptual part and it would then be followed by a constructional part. It was therefore a perceptuo-constructional-cross test, so it was called P.C.C. The original bits of wood were each $2\frac{1}{2}$ inches long. This was considered to be too short for easy manipulation, especially with younger children, so the pieces were enlarged to twice their original size (i.e. to 5 inches), and the thickness was correspondingly increased. The wood was then painted signal red.

**Introduction.**

E. spreads out the 6 bits of wood before S. and says - "Here we have 6 bits (or pieces) of wood. If you look at them carefully, you will see that they are all the same length". (E. pushed the pieces close together to demonstrate) "and they are all the same thickness"./
thickness". (E. turns over one after another to show that although they are differently carved, the actual thickness is the same in all). "The only difference is that they are differently cut" (or carved). (E. invites S. to handle the bits).

This introduction is designed merely to allow the subject to get settled down and acquainted with the test material. A specially observant subject, may however, notice that the pieces, being of equal length or thickness, have possibilities of being fitted together into a symmetrical shape.

The test consists of 10 steps; the first 4 being solely a matter of matching the correct pieces of wood with the diagrams. From first to last steps, however, the test follows a logical progression. The subject is made familiar with the pieces in order of their construction, so that it is possible for anyone who realises this, to learn by experience. Actually the writer is of the opinion that with child subjects, at any rate, this realisation scarcely ever occurs, any more than it does with the Passalong test. The child appears to treat each step as complete in itself. However, if an occasional subject does turn up, sufficiently well poised in the test situation and to whom such a progression within the test does occur, then the connection between the steps does exist for him to utilize. In any case, in the later stages of P.C.C. the development is partly brought to his notice by leaving some of the diagrams in front of him. Without this, the test was found to be impossibly difficult. The more unintelligent subjects, especially the younger ones, did not notice the development even then.
Diagram I is placed flat on the table in front of S. as shown above. The actual pieces are placed below, but reversed in order, and turned upside down.

Instructions:
"You see these 2 bits of wood? Place them so that they look exactly as they do in the drawing and in the same order. (If S. hesitates or is uncertain - "Yes, just turn them around till they look exactly as they do in the drawing").

Time Limit 2 minutes. Score 2. (Deduct 1 for every minute or part thereof more than 1 minute).
Diagram II is placed flat on the table in front of S. as shown above. The actual pieces are placed below, but the pieces turned upside down and the 2 inside pieces moved to the outside, so that the order becomes 2 1 4 3.

Instructions.
"Now place those bits of wood so that they look exactly as they do in the drawings, and in the same order".

Time Limit 3 minutes. Score 3. (Deduct 1 for every minute or part thereof more than 1 minute).
Diagram III is placed in front of S. this time standing upright (against any sort of support, preferably the box for the pieces). Actually this is the way the pieces look when they are to be constructed, so that the possibility of learning by experience can occur. The pieces themselves are placed flat, below the diagram, reversed in order and upside down.

**Instructions.**

"Now place those bits of wood so that they look exactly as they do in the drawings, not only in the right order, but this time in the correct position." (If S. does not stand the 1st piece upright, say once only "now, is that exactly the same position?")

**Time Limit** 3 minutes. **Score 4.** (Deduct 1 for every 30 seconds or part thereof more than 1 minute).
Diagram IV is placed in front of S, standing upright. The pieces are laid flat below, upside down, with the two inside pieces on the outside, i.e. in order 2143.

Instructions.
"Now place those bits of wood, in the order, and in the position in which you see them in the drawings".

Time Limit 3 minutes. Score 4. (Deduct 1 for every 30 seconds or part thereof more than 1 minute).
Diagram V is placed, standing upright, in front of S. The pieces are laid flat, this time in the correct orders, that is, the standing-up piece, which is actually the first piece (regarding the test as a whole), is placed on the left, and the horizontal piece is placed on the right hand.

Instructions.
"Now fit (or put) those two pieces together to look as they do in the diagram, and in the same position". (If S. having successfully put the pieces into position, still holds on to them, demonstrate that the structure will stand quite well unsupported).

Time Limit 3 minutes. Score 5. (Deduct 1 for every 30 seconds or part thereof more than 1 minute).
Diagram VI is placed, standing upright, in front of S. (This is actually the same as the previous step, but the pieces are now made to face in the appropriate dimensions for completing the test). The pieces are again placed flat on the table, below the diagram, again in the correct order. This diagram is left standing, along with the next one.

Instructions.
"Now fit the pieces together to look like that Diagram".

Time Limit 5 minutes. Score 3. (Deduct 1 for every 30 seconds or part thereof more than 1 minute).
Diagram VII is placed, standing upright, opposite S. (Diagram VI is still there). The 3 pieces are again flat, and in the correct order.

Instructions.
"I'm going to leave this other drawing in front of you. Here are the pieces, and the third one for this drawing." (No other indication to be given that one drawing follows exactly from the other, with one additional piece).

Time Limit 4 minutes. Score 6. (Deduct 1 for every 30 seconds or part thereof more than 2 minutes).
Diagram IX is placed upright opposite S, with VI, VII and VIII remaining in the correct order. The 5 pieces are placed flat, in correct order, before S.

Instructions.
"Here is the next drawing, and here are the bits of wood you need".

Time Limit 5 minutes. Score 7. (Deduct 1 for every 30 seconds or part thereof more than 3 minutes).
Diagram X placed upright, along with VI, VII, VIII and IX.

Instructions.
"Now this is the last drawing. And here are the bits of wood".
(Place in correct order.

Time Limit 5 minutes. Score 7. (Deduct 1 for every 30 seconds or part thereof more than 5 minutes).

Possible Total: 50.

Note: To enable the administrator to pick out the required pieces of wood quickly and without fumbling, the code letters (1 + T O X

(1 2 3 4 5 are carried in the top corner of each bit of wood. The 6th bit, being uncarved, requires no letter. Administrator will need some practice with the test, before being able to remember the pieces of wood without the code indication.
Try-out of Test.

P.C.C. was given by the investigator to an unselected group of 103 pupils of Moray House Higher Grade Demonstration School. (1) The age range was 12 - 15. Norms will be established when a few hundred more individuals have been tested. In the meantime the raw scores have been sufficient to calculate the correlation coefficients of P.C.C. with a battery of other tests.

The distribution of scores is given on the opposite page.

The mean score is 21.6; \(\sigma_m = .98\) and the standard deviation is 10; \(\sigma_r = .70\).

It will be seen that the distribution is slightly positively skewed. This is quite acceptable, however, for it was hoped that the test would reach adult level; and head-room is therefore left for highly Superior Scores. The writer gave the test to a group of 50 children of superior I.Q. It will be seen from the second graph that some of those scores do fall at the upper end.

\(r\) with Intelligence = .438.

Unfortunately only 57 of the 103 children tested were found to have an I.Q. recorded for them in the school records. The intelligence tests used were Moray House Intelligence Tests, from 2-4 years old, and these would probably have been slightly more verbal than the current ones, especially the 4-years'-old one. (2)

The \(r\) is significant. In Fisher's tables an \(r\) of .34 is significant at the 1% level with 57 cases.

It is hoped that the test would measure mainly \(g\) and \(k\). The actual \(r\) with intelligence test scores is therefore approximately about what might have been predicted.

(1) The group was unselected for intelligence, but might possibly be regarded as selected socially, relative to the general population.

(2) Statistical Tables for Research Workers.
$r_{\text{with Mathematics}} = .58$ (Fisher's $Z$ value .67).

It occurred to the writer during the testing of the group that the differences in ability to do the test were so marked that she wondered what relationship there might be between it and mathematics, geometry in particular. (Perhaps it would have been more instructive if the $r$ between geometry alone and P.C.C. had been calculated.) The test was given to all 3 years of the Higher grade school and the $r$s with mathematics were calculated separately for each class. The average $r$ is .58, but better interpreted in terms of Fisher's $Z$ value it becomes .67.

**Sex difference.**

Since significant sex differences have been observed in performance tests, it was interesting to see how new boys and girls fared on P.C.C.

\[
\begin{align*}
M \text{ (boys)} & = 21.605 \ (M = 55) \\
M \text{ (girls)} & = 18.156 \ (M = 38) \\
D & = 3.449 \\
st & = 2.1968
\end{align*}
\]

It is a pity that there were almost twice as many girls as boys in the sample. (One would have been justified in anticipating that the numbers would have been more equal in what was the whole of the Higher grade school). It does not quite reach significance but there is a very decided trend in favour of the boys. It will be interesting to see if this value increases as we gather more cases.

* We use Fisher's $t$ instead of $\frac{D}{s}$
The position of the Visual Perception of Space in P.C.C.

As we have already pointed out, the first part of P.C.C. is purely perceptual while the second part is constructional. Steps I to IV are a more or less simple 3-dimensional situation. Yet the types of response to this situation were so varied that the writer was very much impressed by what appeared to be three different ways of perceiving depth on the part of the subjects — not directly unconnected ways, but as points on a "scale" of perception.

The range began from those subjects who appeared instantaneously to see the diagrams as clearly distinguished drawings in 3 dimensions, each one an easily recognizable 3-dimensional gestalt. It was the readiness of such subjects to match the appropriate pieces of wood, that was such a distinguishing characteristic. All they needed to do was to glance at the diagram and then with unfailing promptitude to pick out the correct piece. This way of seeing seemed to be of the immediate (1) gestalt type where relations are regarded as a secondary product of analysis of the total form, this itself being primary and primitive, inherent in the process of sensory reception and not dependent fundamentally on learning or experience. Even the modification of gestalt theory to account for the facts of form blindness in brain injuries, makes no difference introspectively between sensation and perception. 3-dimensional perception, then, ranged from this type of response to that where central factors of set and meaning were operating, in which the subject did/

did not see the diagrams as shaped in a certain way, but saw that they were shaped in a certain way. That is, the subject saw not immediate forms, but things, and subject's effort in sense perception was to know the objective situation.

At the other end of the scale were those subjects who saw neither immediately in the first way, nor gradually in the second way, but in a slow laboured fashion, using as it were 2-dimensional means for a 3-dimensional end. In other words they saw the diagrams quite flat. This was most obvious in one or two cases where the boy or girls traced out the form on the diagram with a fore finger, and then fingered it out on the bit of wood. If allied with low intelligence this type of perception was rather painful to watch. The impression of the writers that the diagrams were confusing because of their flatness was confirmed by the subjects' introspection.

This leads the writer to put forward the suggestion that a psychologist's theory of form perception will depend to a great extent on his own type of perception. After all, as H.G. Wells has put it, "One can only be one's own rabbit". Also, the writer often found that she herself failed to see some of the Gestalt figures in the way in which the Gestalt psychologists assumed that they were seen. Nor did she find that some of their principles such as the Law of Pragnanz, held so universally as they were supposed to do. She found that she quite frequently saw a figure as a-symmetrical when according to the Gestalt theory an "internal force" ought to have operated in favour of her seeing a more simple/
simple, symmetrical configuration. In fact, she believes, that these individual differences in perception of form account for the failure on some occasions to get the same results as the Gestalt psychologists, when other workers have repeated their experiments. There appears to be in some individuals a 'constitutional' tendency not to see things as things, but to see that they are things, that is, to use inference rather than intuition. The Gestalt theory generally, is an intuitive theory. A Gestalt psychologist would see a field as a field. The other type of 'perceiver' would see that it is a field. (In this connection also, the writer is of the opinion that Spearman's two steps in analysing a pattern, namely, the apprehension of the constituent fundamentals followed by the education of relations, are not to be regarded as inevitably present. The first may exist by itself. We therefore do not accept Kousay's conclusion (following Spearman) that this, in all its aspects, must involve the retentivity observed in all group factors by the Spearman School.)

In a general discussion of "wholes", also, this immediacy of the Gestalt type of perception is illustrated by a remark of Krüger of Leipzig "What wholeness is we learn immediately and with an unquestionable certainty in our experience".

There seems to be a corresponding broad division applicable in the whole field of philosophy. From the earliest times many philosophers have objected to analysis. They have maintained that analysis is falsification, that a whole does not really consist of parts suitably arranged, and that, if we mention any part/
part singly, the act of isolation so alters it, that what we have mentioned is not what is an organic part of the whole. The extreme of this point of view leads the holder to maintain that there is knowledge not expressible in words. Words are nevertheless made use of to tell us what this knowledge is! The holders of this viewpoint include the mystics, Bergson, and Wittgenstein, and we find it in certain aspects of Hegel and Bradley.

The principle of atomicity represents the opposite extreme. It may be said to forbid synthesis. Linguistically it forbids the giving of proper names to complex wholes, at any rate when they are recognised to be complex.

Somewhere between these two points of view stands Bertrand Russell. In discussing visual perception (1) he says that he personally does not see a black dog on a background of snow as a whole; he tends to see the black dog. Similarly what he notices after the rain is a rainbow, not a rainbow plus some sort of ground. He admits the usefulness and the frequent applicability of seeing "in wholes", but he refuses to carry it to an extreme. He says (2) "It is customary nowadays to dismiss contemptuously the atomic view of sensation as it appears in Hume and his followers. We are told that the sensible world is a continuous flux, in which divisions are unreal, the work of the mind purely/

(1) An inquiry into Meaning and Truth: George Allen and Unwin. 1940.
purely conceptual and so on. This is said as something obvious, for which only a stupid man would demand evidence. Now the word "sensation" or "sensible", as is often pointed out, stands for something hypothetical - broadly speaking, for what could be noticed without change in the environment or the sense-organs. What is not hypothetical is what is noticed, not what could be noticed: and what is noticed has, I maintain, just that atomicity and discreteness which the critics of Hume reject. They do not, as empiricists should, start from data, but from a world that they have inferred from data but use to discredit the kind of thing that can be a datum. In theory of knowledge what is fundamental is noticing, not sensation.

I shall take for granted, henceforth, that we can, within a perceived whole, perceive parts as inter-related".

We give this quotation from Russell to illustrate what we might call a possible "middle" point of view on our visual perception scale. He personally rejects the monist and atomic extremes, but although he sees the force of Gestalt-like perception in certain circumstances, he yet maintains the more atomic view as appropriate in other circumstances. As a "general" attitude he would adopt the inferential one.

Summing up our view of visual perception of space, then, with specific reference to the first four diagrams of P.C.C., we may represent it roughly thus:

<table>
<thead>
<tr>
<th>Pure Gestalt Perception</th>
<th>Analytic Perception</th>
<th>Mosaic Perception</th>
</tr>
</thead>
</table>

*No doubt Russell is here deliberately laying himself open to the charge of "bifurcating the universe", to use Whitehead's well-known phrase, since he is plainly separating thing from environment.*
We have no exact measure, but we should say from observation of the individuals doing P.C.C. that the distribution tended roughly towards the normal curve of distribution.

In the pure gestalt type, it seemed as if each 3-dimensional figure was a "pure" datum. The whole and the parts were equally present in consciousness and the immediacy of the perception enabled the subject readily to match up the appropriate piece of wood with each diagram.

The mid-type could be described thus. The perceptual datum at first may appear as a somewhat vague whole, but subject gradually arrives at an enumeration of the inter-related parts; but in such a case it may be said that the datum changes as a result of attention. This is certainly true, in the case of a visual datum which we observe first carelessly and then attentively. Attention in such a case involves changes in the eyes, which change the visual object. It may be said that all analysing is of this sort, and that the whole whose parts are known is never identical with the previous whole which was perceived vaguely.

At the other end of the scale, is what we might call the completely mosaic type. There is no reference to a whole. The perceiver sees but a maze of parts, as for example in P.C.C. a small number of subjects appeared to be seeing only flat figures with bits in the middle sticking out, or in, as the case might be. Generalizing for a moment, - it is a case of not seeing the wood because of all the bits of trees. In an unintelligent subject, as we have already observed, this type of perception may lead to rather chaotic results, at least in the visual perception of types of 3-dimensional diagrams. Harkening back to the girl in the factory, /
factory, whom we mentioned as being completely unable to see the cubes, it is possible that this type of form-blind or mosaic perception applies to her.

Since we have been discussing types of form perception and their possible influence on the type of philosophy, accepted or separately evolved by a thinker, we should like to draw attention to one more point concerning the influence of the "personal equation" in determining fundamental attitudes.

The pre-occupation of the Gestalt psychologists with "form" is interesting because it shows to what extent the "aesthetic" plays a part in this theory. Since Aristotle we have seen how form has attracted the attention of observers and it is therefore not new in itself. Perception of it supplies the bulk of artistic appreciation. Form has been cultivated along special lines e.g. the melodic form by musicians, spatial form by artists, and literary form by authors. It has been rather neglected by psychologists. It is characteristic of the aesthetic attitude, which seeks to comprehend the intrinsic harmony of any object that is the centre of attraction. The object may be as trivial as an ornament or as substantial as a human being. But it is the singleness and the symmetry of the structure which interests the aesthetic person. When highly developed it may to a certain extent offset limitations of experience and intelligence, but when it is combined with these qualifications it enormously enhances a judge's skill. The Gestalt psychologists appear, to the present writer, to combine the aesthetic attitude with an intelligent theoretical standpoint, although the intuitive, aesthetic side predominates.
It is interesting to note, regarding the supposed opposition between the theoretical, "cognitive" seeking after "Truth" attitude, and the aesthetic outlook, which tends to believe with Mencken that "to make a thing charming is a million times more important than to make it true", that in the last analysis subjective factors play an enormous part in the creation of scientific theories, just as we believe they do in theories of perception. Take the attempts of Einstein and Eddington to reduce the laws of electromagnetism to geometry. Einstein dislikes Eddington's theory, although he sees no way of disproving it, while Eddington has made the remark regarding Einstein's theory that it is a "matter of taste". J.W.M. Sullivan (1) has given an admirable account of how in certain cases, when ordinary "objective" criteria fail, the final criterion may be an aesthetic one. Other examples of this conflict between psychological types have been given by Poincaré who divided mathematicians into psychological types, and showing that a kind of reasoning which would convince one type would never convince another. These differences he claimed to be fundamental and to play a great part in the actual construction of science.

The greatest drawback of the Gestalt Theorie, in spite of Koffka's discussion of the Ego, is the discounting of importance of the perceivers. As Stern said, "Keine Gestalt ohne Gestalter".


(2) Compare also, not as in opposition, but from the point of view of individual differences, the different approaches to factor analysis - the algebraic approach and the geometrical approach.

Gestalter". Both he and Rignano objected that even the significance of the phenomenal Gestalt is invariably its relevance to the person in his intricate process of adaptation to the complexity of the world. Similarly the essence of space and time is their personal relevance.

We have gone into form perception, and individual theories, in slight detail for it appears to the writer that with some of the thinkers we have mentioned, it becomes wholly absurd to account for fundamental differences in their attitudes and theories in terms of anything so simple, for example, as I.Q. Above a certain I.Q. level, numerical additions to it become meaningless. We must seek other causes, such as individual exigencies of temperament and special abilities, or special factors and clusters of factors. Such inter-acting influences, are of course, equally important with all levels of I.Q. It is for this reason, therefore, that we venture to suggest that differences in the visual perception of space, if pursued in all their ramifications, are of a fundamental importance.

The Space Factor.

So far we have discussed only the first stages of P.C.C., that is the steps involving visual perception alone. Steps V to V however involve some sort of constructional ability, as well; or at least some sort of ability to imagine the pieces as they would be, if properly constructed.

It was occasionally found that a subject who showed marked ability to place the pieces correctly in step I - IV, failed when it came to the constructive stage. Sometimes such individuals, on/
on being questioned, said that they were good at art in school; and perhaps in their case, speedy perception of form was an important factor. Such "good" space perception is, obviously not enough, however, to enable them to continue so competently with the rest of the test. If it were outstandingly good, it would undoubtedly help, but the use of imagery appears to play the most important part with steps $\bar{Y}$ to $\bar{X}$, if the subject does have competent visual spatial perception. Individuals of the latter type appear to do the test more rapidly, but there did seem to be some subjects who, possibly having low "k" (1) utilised some other means at their disposal to do the latter portion of the test, making what seemed like a verbal formulation first. That is, they, verbally described to themselves each bit of wood, what was missing, and what had to be filled up. This method of course lengthened the time taken to do the test. Sometimes subjects who appeared to be working it out thus, exceeded the time limit. They were not credited with a score, but they were allowed to go on for a time to see if they could actually finish. Some did; some did not. We may say that such individuals had low k, or else we may say that their actual visual spatial perception was of our middle type and tending to the mosaic end. But "good" visual perception of form is helpful with the second part of the test too. It is not just a matter of being able to use spatial images and spatially imagine them into the correct position. There must be good, plain, 3-dimensional "seeing" as well, and when it is good it has this immediacy which is so markedly absent with our middle/

middle analytic type. It has the spontaneity of something innate, although it is the writer's opinion that some improvement in 3-dimensional seeing may be effected, given a reasonable level of intelligence and awareness of what is necessary. It would appear, when we talk of "improvement" that we are assuming that our middle type is deficient, and that therefore we ought not to talk of type, but of degree of k factor. This however, does not necessarily follow from our remark. We are merely saying that when dealing with visual spatial material under certain conditions, e.g. in school geometry, it would probably be a good idea to determine the type of visual spatial perception and help it specifically for geometry without necessarily attempting to change the fundamental type of perception, which may possibly tend more to the "inferential" than to the gestalt. A simple diagnostic test might be the cube-counting, as used in the American Army Beta test. Differentiating the extremes of our types we might possibly find that the "Gestalt" perceivers "saw" well and rapidly, that the mosaic type at the extreme end could not "see" at all, and that shading into either end, the analytic type might gain as high a score as the Gestalt type, but might take a longer time or else do less well, (by taking too long a time.) All types, of course, would "perform" relative to their level of intelligence. There is good evidence that "g" plays a very important part in all spatial tests.

Next, we must take note of the sex difference we have observed. As we shall see, in discussing our battery of performance tests, the sex difference was apparent only in tests involving obvious spatial relations and not in the Healy Picture Completion II for example.
example. The first and most readily apparent reason for the sex difference is of course the greater familiarity of boys with this type of material than girls. We believe that this, plus a spatial factor, plus an individual temperamental attitude, are all part of Alexander's performance factor, and that a simple "performance factor" does not exist.

On the other hand, there is the possibility of the sex difference in *k*-factor being innate. We would rather discount this, although the possibility as such must be noted. It has been noticed recently in some test results that girls tend to score higher on verbal items than boys. We cannot help wondering to what extent this may be some "compensation" for inferiority with material involving spatial relationships. It is perhaps not too far-fetched to suppose that centuries of the attitude, that, what is concerned with *overt* action and the *concrete*, "real" environment, is a male 'preserve' only, may have led to verbal predominance or preference on the female side. It is scarcely to be wondered at, that women are supposed to have placed more reliance on intuition than men. There was little else for them to have relied on, since restricted environment curtailed the objective criteria which were essential bases for judgment and a higher level of thought generally.

Finally,!

(1) We are using "temperamental" here in the wide sense, as is current in English psychological usage. Although fundamentally we regard Temperament as innately determined, we prefer to give it a wide connotation, making as little use as possible of the term Personality.

(2) Intelligence, Concrete and Abstract. B.J.P. Mon.Suppl.12-1935.

(3) The spatial factor may include, we believe, a cluster of factors, one being *k* factor which is confined specifically to visual space perception. (There may also be different rates of maturation which must also be taken into account.)

(4) It has been found also that Cox's "m" factor disappears when knowledge of mechanical principles is not present.
Finally, it would appear to be useful if some sort of comprehensive investigation of the whole field of spatial perception could be undertaken.

Koussy’s K-factor does appear to account adequately for the tests which he himself used. On the basis of the introspections of adults and children he shows that for those tests involving K factor "it is necessary to obtain, manipulate and utilise visual spatial imagery". We should like to point out, however, that, although as in some of Koussy’s tests subjects reported inability to complete items which involved the choosing, for example, of a rotated shape out of several such shapes, because of their inability to visualise the spatial material, we must not assume that visual imagery, as such, was poor in these subjects. It was specifically visual imagery of shapes. To illustrate what we mean, we may mention the introspection of two adult subjects with regard to the National Institute Space Perception test which Slater evolved to measure k-factor. It consists of just this matter of rotating and general "manipulation" of shapes which we have mentioned. One mathematician did agree that he could only solve the items by imagining the shapes as they would look when turned around. Another adult subject, with less efficient k-factor, was not found to be deficient in visual imagery generally, but only in visual spatial imagery. This subject, for instance, could very easily get a clear visual picture in both black and white and colour of somebody or other standing in back-view, and could visualise at will and describe very accurately many other varieties of objects seen as visual images. (1) It would be useful/

(1) This subject also reported unwillingness to deal with the shapes, and introspection indicated emotional "blocking".
useful if one could have the introspections of two artists, in
further investigating this matter; the one would be the type of
artist who must always have a model, because of inability to
retain images, while the other would be one who does not need to
draw "from life" because he can rely on competent visual imagery.
It would be interesting to see the relative predominance of k-
factor in those two subjects, since the difference we have
mentioned in their case was not of spatial visual imagery, but of
visual imagery generally.

A further observation about the field of space perception.
There are many problems which need solution. For instance
are we to account for people, suffering from brain injuries, who
are so grossly disorientated that they are unable to walk through
a door, but bump into the wall instead?

And what again, is the essential element in orientation
tests? How shall we account for the immediacy of the reply
of a person, given for example the kind of question which involves
their own position and the position of the sun. Certainly the
simplest types of this question in which only one's own position
is concerned can be correctly answered by imagining the map of
Britain, or one's own position facing a known direction in a
town, but only in simple orientation tests is this a cue. What
of the more complicated ones? Professor Drever discounts the
importance of imagery here, and is of the opinion that the
important factor is of a kinaesthetic nature. He thus takes
up the 'nativistic' position, and would also extend the importance
of kinaesthesis to the wider field of space perception.
Our conclusion, is therefore that although K-factor appears in K-factor tests, there is no group space factor, but a cluster of different factors or types of space perception, of which K is but one. A really comprehensive investigation using tests of many aspects of space perception is indicated.
CHAPTER IV. THE INTER-ACTION WITH TEMPERAMENT OF INTELLIGENCE AND SPECIAL ABILITIES.

In the first place we had perhaps better say what we mean by intelligence, though we shall scarcely attempt a straightforward definition!

By intelligence we do not only mean "what is measured by intelligence tests", for so little really comprehensive work has been done on finding out what individual items of such tests are testing, to take one aspect of the question alone, that it seems unwise to make ambitious assumptions. What the present writer does understand by intelligence is not only what is measured by intelligence tests, nor yet what is measured by what are called "g" tests, but something which seems to indicate the general ability to deal insightfully with relations.

Of one thing, however, the present writer feels fairly certain, namely that the intelligence test situation measures nothing so "pure" as intelligence, and that factors such as temperament and special abilities are playing proportionate parts, according to the extent to which they are evoked by the general situation. This should be a psychological platitude, but it is remarkable how often, and to what degree, it is totally neglected.

When Binet in 1905 proposed to erect a definite scale by which man's intelligence could be measured according to an exact standard, a really practical analysis of human personality was/
was initiated. Intelligence testing is now firmly established in our educational framework, but although an enormous amount of knowledge is yet to be gained, we must yet admit that one of the factors of personality is being measured in a reasonably satisfactory manner, relative to our knowledge.

We all know, however, that having assessed a child's I.Q. we are very far from having an analysis of his make-up that will enable us to understand his temperament, to predict his behaviour, or to decide his later fitness for a position of responsibility. Experts in vocational guidance, child guidance and education all stress the significance of factors other than the intellectual. The great drawback is lack of reliable tests, and lack, also, of decision as to what the variables are which we must measure.

The question of the influence of instinct on the development of intelligence requires a volume, or volumes by itself. Take a timid, fearful child for example. He would be continually on the defensive, always paying attention to innumerable details in his environment, such as a frown on somebody's face, - details that would escape a more fearless companion, with more assurance in his own ability to deal with such contingencies. How would such a child react in an intelligence test situation? Probably well enough, and in proportion to his intellectual endowment, if there are not too rigid time limits in force, but an individual of this type may tend to have a kind of generalized temperamental attitude which would influence his intellectual "tempo". The score of such an individual compared with that of one who is temperamentally on the aggressive and less pre-occupied/
pre-occupied with detail, is mainly a score, relative to those particular conditions. This is the type of individual variation in temperament which is obscured in group testing, though it has been suggested on the other hand, that the group test situation often is beneficial to the timorous, self-conscious individual, in that he thus loses his identity. There may be indications, however, of temperamental attitude in the scatter of individual scores in group tests. Some very interesting surveys of this have been published.

Constancy of I.Q. is another problem which has drawn out conflicting results. Data have been published both in support of such constancy and equally contesting it.

One of the general questions which arises in intelligence test situations is the influence of excessive orexis. It may be a hindrance on cognition and cause flurry. Philp(1) observed this in some of his subjects when they were solving verbal puzzles. In studying 'frustrated conative effort' he found the correlation between puzzle-solving ability and intelligence to be .20. Philp decided that puzzle-solving ability may be due to certain character or to certain perseverative qualities. He put forward the suggestion that puzzle-solving ability often accompanies the absence of emotionality.

It must not be assumed that the factors which we are mentioning as influencing intelligence test results, imply that we are adopting an anti-intelligence test attitude. That would be absurd. But we certainly believe that such factors are of vast influence, and that all degrees of such influence may operate/

operate.

When we decided to use a battery of performance tests, we realised that with the higher I.Q. levels there may be a correspondingly complex personality integration. We thought, however, that we might perhaps get around this difficulty by the fact that our aim would be disguised, and that if such subjects imagined that the point of such tests (which we made as difficult as possible) was to see how well they could complete them, then if there were any generalized temperament traits, those would be revealed in the effort involved in achievement.

R. Dellaert from an examination of unstable and inferior subjects showed that their inferiority was more marked in tests of verbal type, and considered performance tests important for the assessment of social aptitudes.

Jastak investigated a large number of American school children and discovered evidence of differences in "level" between vocabulary tests and tests requiring some degree of performance ability. His results indicate that grossly maladjusted children tend to obtain much higher scores in the vocabulary than in the performance tests, while normal subjects tend to have similar scores for both types. In cases of mild maladjustment the scores are somewhat irregular, but tend to be higher for performance than for vocabulary tests. These differences, Jastak holds, are due to variations in certain non-cognitive/

(1) "L'Intelligence des anormaux du caractere." L'annee Psychol., 1934, XXXIV. P. 200-216.

(2) Variability of Psychometric Performances in Mental Diagnosis. New York, 1934.
non-cognitive qualities which may be classed under the heading of "mental control" and depend "on those forces which control the ways of intellect rather than on its level". Jastak suggests further that the introduction of the factor of complexity of response, as distinct from intellectual difficulty imposes a need for inhibition and control which poorly integrated personalities cannot meet, thus explaining their relatively low scores on performance tests. Later, Jastak investigated mental patients in a similar way, demonstrating a difference between the score patterns of psychotic and non-psychotic subjects.

A study of adult morons was carried out by Earl (1) who found that those who showed a marked preponderance of verbal ability tend to be unstable, with a neurotic preference for symbols rather than for realities. Where verbal scores were markedly lower than performance, a "verbal neurosis" was assumed, which was believed to reflect early educational difficulties.

With young adult morons, Earl found that the orthodox technique was not satisfactory. A mental age of 9, for instance, may be present in a socially adequate individual, while another of mental age between 11 and 12 may be defective by a social criterion. Earl suggests that at any rate with the subjects with whom he came in contact, social adequacy, rather than I.Q. was the essential point - in other words the subject's general ability to survive in society.

The Mental Age of 9 is held to be the level of mental defect/

defect on the Terman revision of the Binet scale. The Performance level, however, should be higher, for verbal symbolic ability is phylogenetically a late acquirement so that in sub-normal cases we may expect more of a defect here than in the spatial-objective field. Also there may be some improvement in performance scores in moron children which is related in part to maturation, without reference to increased intelligence as measured by the Terman scale.

The Performance limit for defect is 11 years. This, however, produces a paradoxical situation in which a subject, whose performance range is 9-11 years on three sound tests is quite likely to adjust well in society if his Binet age is below 9, but he is extremely unlikely to do so if it is 11 years or more.

This may be partly explained by what is a fundamental difference between the two test methods. The Terman is largely a test of what Thurstone calls the "product of intelligence", while performance tests measure intelligence in operation. The Terman revision of the Binet and such like linguistic tests call for very little persistence, or what Alexander\(^{(1)}\) calls the "x" factor. No very great effort is needed to "tell what is foolish" about certain situations, or to explain proverbs. Either the subject can do that sort of thing, or he cannot. Either he can think in abstract terms, or he cannot. But in performance tests, on the other hand, especially in lengthy performance tests, he must persist and retain control in the face/

face of increasing difficulty. And what is also important is that the subject must inevitably become aware of the mistakes he makes to a far greater extent than in the Terman. One block out of place will prevent subject from matching up exactly his block design with that of the diagram. He sees this, and the strain remains. The test administrator cannot say so readily that he can pass on to the next question.

In life and in clinical practice intelligence is expressed for most people in terms of behaviour value, and not in abstract thought. Thought, biologically regarded, is incomplete behaviour. It occurs between the impulse and the overt act. The abnormal subject, the neurotic and the unstable, nearly always prefer thought (maybe fantastic thought) to objective reality; incomplete behaviour to complete. As a result the subject's interest is determined in favour of the verbal-symbolic field. As Earl has said (1) "It is this fact which accounts for the neurotic intelligentsia at one end of the scale and the unstable defective at the other." We must therefore remember that when we compare the results of verbal and performance tests in a neurotic subject we must remember that the first is testing the field in which his special interests lie, while the second is imposing upon him the unwelcome emotional strains of decision and of complete behaviour. In the defective, Earl says that any difference is even more significant /

significant, for whilst his interests may be directed by his emotional abnormality into the verbal-symbolic field, his achievements in that field are limited by his low intelligence - his poverty of concept formation. If, in spite of this, the defective has a performance age below his Binet age the evidence of emotional abnormality is all the stronger. In clinical practice most defectives have a performance age definitely above the Binet age and unless this is so, the prognosis for resocialization is very poor.

It follows, therefore, that a large discrepancy between a performance test result and a verbal test I.Q. ought to have considerable significance for the entire personality. It will give some indication to begin with, of whether or not the mind is, potentially, of the all-round type, (relative to the limit of difficulty of the tests, of course). It ought to reveal, also, what we might call the "patchy" mind, which may have a high special ability, such as a verbal one, which although special, may yet be used to a high creative degree by the individual. Finally, if there has been a "temperamental" link with disability in spatial relationships, or with overt action generally, this ought certainly to be revealed.

As a result of wide experimenting with tests, the so-called verbal factor or "v" has been psychologically established. It was gradually realised that the earliest intelligence tests tended to be highly loaded with this special verbal ability. We can all think of examples of individuals who, endowed with a certain verbal fluency and coupled with what we might call high/
high positive temperament traits, have seemed to exceed in achievement what we might have predicted for them on the basis of their general ability. Alexander(1) demonstrated to what extent "v" was operating in the battery of tests which he gave to different groups of subjects, and also pointed out for which school subjects, as a result of analysis, "v" seemed to be important.

Head(2) in his study of aphasia, gave distinct and incontestable evidence of an area of the cortex which appeared to be specific to speech, for when it was damaged, a corresponding deterioration or loss of the power of speech occurred. Head therefore made the evidence for the existence of a special verbal disability quite clear, so that we may therefore assume the "verbal factor."

Not so much attention was paid to items in tests involving spatial relationships, although now that it is recognised that some such factor or factors exist, tests have been evolved such as those of the National Institute, which present, in pencil and paper form, an economical and more exactly standardised way, what was the essential element or elements in some types of performance tests. Alexander, in his survey, isolated a factor which he called F, which, along with g, made up the performance factor. This, in meaning at any rate, corresponds to what we mean by a special ability to deal with spatial relationships. Coats mechanical factor "m" was similarly designed to cover this special factor, although he stated that it/

it was not meant to cover performance tests. More recently Slater has devised tests, which we have already mentioned, to measure visual spatial perception, on the basis of Koussy's findings in this field. Summing up, then, we may assume that psychologically, evidence exists, strong enough to allow us to accept the 'reality' of some such factor which operates independently of the verbal factor and to some extent independently of intelligence.

Now, what may be said of the neurological evidence in this field? We find that Kleist (1), in dealing with brain injuries observed after the last war, gave evidence of what he called "constructional apraxia". He found that patients were unable to copy simple designs, build with bricks, draw or write. Their capacity to deal intelligently with spatial-objective construction was quite lost. Thus we find that there are good grounds for supposing that an area, which when damaged, leads to this disability, may be held to correspond to the area equivalent to that for speech.

Neuro-psychiatric evidence is therefore in favour of both a special verbal ability and a special spatial-objective ability. There was actual neurological evidence, too, 15 years before Alexander or Stephenson cited psychological evidence for the existence of such factors.

So far, however, it has been impossible to say how much those special abilities depend upon psychological factors such as interest, or type of imagery, and how much on anatomical or physiological/

(1) Kriegsverletzungen des Gehirns; usw., Leipzig 1934.
physiological superiority. Also it is possible that there may be different maturational rates. We know that visual space perception, studied genetically, has given us evidence of the effect of maturation, but what would be equally important would be an investigation into individual rates of maturation. Other factors might also be included, so that we could see if a special ability such as the verbal one, also shows individual rates of development. The emotional and temperamental difficulties which we may observe interacting with intelligence development, may possibly, therefore, be among the symptoms of unequal rates of maturation of special abilities.

It is likewise important to notice that the twists thus given to temperamental tendencies may determine the use that will be made of general intelligence, although it does not determine the degree of intelligence.

More than ever, it becomes apparent, then, that greater efforts must be made to make a "total" approach to the investigation of each child, at the 11+ age for example, and not confine one's measure to tests of intelligence and school attainments by themselves. It has been shown in several surveys that in one individual the maturity level of temperament and intellect may be attained at widely different times. In fact it has been claimed that a certain fluidity of temperament is better for the development of intelligence than a too early temperamental "set". Downey (2) for instance concluded that:

(2) The Will-Temperament and Its Testing, 1924.
that a high I.Q. plus temperamental immaturity was more favourable for the future achievement of a child than high I.Q. plus great temperamental maturity.

As regards the influence of imagery in testing material R.O. Filter(1) made an experimental investigation of character traits in which he used the Downey(2) Will-Temperamental tests, and criticised them on the grounds that the objective material he utilised involved a visual factor, and that visual endowment as a special capacity may have complicated the situation. We are thus led to ask the question whether or not specific capacities themselves may mask temperamental reactions.

Downey came to certain conclusions on the basis of Filter's criticism which seem to the present writer very important. A temperamental pattern, she decided, carries with it implications concerning the quality of intelligence, although not indicating the level of intelligence. She divided her subjects into four large groups describing incidentally the type of imagery which she believed to be common to each.

The first group were of the slow, deliberate, accurate, highly controlled type. They learned somewhat slowly, particularly when rote memory is involved. In attacking a problem of habit formation they were liable to think it through the first time and take an excessively long time for the initial trial, with rapid gain in speed thereafter. The intelligence of this group was of "the brooding type", an outcome probably of:

(2) Downey herself found also that with young adults, men showed much greater power of inhibition than women. Actually 70% equalled or exceeded the median record of women. She also found that the college men showed a tendency towards a more definitely patterned reaction than the college women.
of strong perseverative tendencies. Imagery, Downey found, was very visual.

The second group were quick-reacting, rapid-fire and explosive types, only slightly interested in detail, with some inadequacy on the side of motor inhibition. On the intellectual side they learned easily. They warmed up quickly, shifted their attention easily and were readily distractible. Given to rapid generalizations, they make use of verbo-motor material and in association tests, show dependence upon verbal habits. Visual images are fragmentary and difficult to hold. Auditory imagery was conspicuously absent.

The third group showed less extreme temperamental patterns, or irregular non-specific patterns. There was low interest in detail. The quality of intelligence was less obvious than with the first and second groups. There was possibly greater versatility with less special efficiency. There was both verbalization and visualisation, but visualisation might show inadequacy or demand special effort. In speed of learning the third group was between the first and the second. Attention was somewhat less readily adjustable, but less distractible, than Group 2. There was greater ease in rote learning than in the first group.

The fourth group showed a generally low temperamental profile. On the intellectual side there was a lack of responsiveness and a general colourlessness of reaction. Intelligence was not necessarily low, but there was a lack of self-confidence, initiative and energy. The members of this group were found to/
to visualise in a limited range.

The part which imagery plays in learning was long ago realised to be of importance. Meumann found that a visual tendency was one characteristic of the slow learner. Angell, on the other hand, found during an experiment that several of his quickest learners were very visual. However, we may express this in rather a different way. It is not a case of saying that visualisers learn slowly, but that those who learn slowly are visualisers. That visual mechanisms have greater inertia than the auditory is shown by the different persistence of sensory after-effects from visual and auditory stimulation. Based on this fact Downey put forward the hypothesis that the retardation of discharge tendencies in the nervous system gives a chance for visual images to develop. One may then expect to find precise and persistent visualization in the case of those individuals who exhibit an unusual amount of nervous energy or "load". The voluntary retardation of motor discharge may also result in visualisation, while fragmentary and evanescent visual images may be reported by individuals who show average "load" or average motor control. Downey claimed that she was thus able to show from her subjects that nervous inertia may result in slow learning and visual preoccupation.

The present writer believes that the whole question of imagery should be re-opened. We have seen how the results of Koussy were based on the "manipulation of visual spatial imagery" recorded introspectively by his subjects. Koussy found also that visualisation proved to be a hindrance in some other/
other tests for subjects who found that too-ready visualisation hampered them in a test involving visual material, but which was actually concerned essentially with deduction of relationships. Considering the importance which we attach, also, to the relative predominance of certain special abilities in an individual's make-up, it should prove fruitful to relate types of ability to types of imagery.

We are very strongly of the opinion, too, that test results which may be of practical importance, such as those at the 11+ age in state-aided schools, should not be obtained from group tests only. Supplementary individual testing, preferably of a performance nature, is an absolute necessity. Any such test is in effect a sample situation from which inferences may be drawn regarding the subject's behaviour in the situations of everyday life. It may be less important, from this point of view, to know the score which a subject makes than to know how he makes it, what attitude he adopts towards the test as a problem, by what method he attempts its solution, and with what degree of consistency he applies that method. His persistence, or lack of it, in situations of difficulty, his reaction to failure, signs of excitability, impulsiveness, or inhibition, and many other subtle variations of conduct, all have a bearing on significant aspects of the personality.

Assessments by means of such clues has of course some very clear disadvantages. It is non-quantitative and in the main subjective, and must be based on rapid interpretation of momentary/
momentary phenomena. The value of the data thus obtained must
depend to a great extent on the knowledge and skill of the
tester in relating the test impressions to the analysis of
the individual's temperament. We shall see whether anything
emerges, which can be observed objectively, in our later survey
of performance test results.
CHAPTER V. DESCRIPTION OF TESTS USED IN THE INVESTIGATION.

A. A Battery of Performance Tests.

As has already been stated, the primary purpose of the writer in using performance tests was a qualitative one, but it was considered important as well to see if some more light could be thrown upon what exactly such tests were testing.

The use of performance tests seemed to be especially valuable from the point of view of observation of temperament, for this appeared at any rate to be a way in which an investigator could start out with as few a priori assumptions as possible, and hope to be able to see some "real" and characteristic reactions emerging from the individual methods of attack. It seemed to be a way also, in which to a certain extent, the investigator might obviate the most likely difficulty to occur, namely the making of general temperamental rules out of the exigencies of one's own temperament.

Next, as we have mentioned earlier, discrepancy between verbal I.Q. and Performance score seemed to us to be of significance temperamentally, in proportion to the size of the discrepancy.

What seemed of especial importance was the fact that some control could be exercised over motivation: not control in any absolute sense, for that would be impossible, but control to the extent that the aim of the tester was disguised. Before each subject began the tests, the writer said a few words/
words to the effect that achievement in those tests would not affect school conditions in any way, but that since the results were being used for research, it would be useful if the tests were done as well as possible. It thus looked as if achievement in the tests were the only thing that mattered, so that if there were any constant temperamental characteristics really typical of each individual, they might be expected to emerge while the subject was unaware that anything more than test achievement was being observed. Motivation has always seemed to the writer, to have wielded such an uncertain influence over the results of temperament tests that it seemed essential somehow to bring it to operate under as normal conditions, (and therefore entirely particular to each individual) as possible. Preliminary stimulation of the competitive spirit, in that subjects are told that results will be posted up, as has been done in several reported investigations, has always appeared to the writer to nullify the results, and to intensify the temporariness of the subject's effort. Surely what we wish to find out is what is characteristic, and therefore fairly constant, in the individual's temperamental make-up.

Finally, the writer is strongly of the opinion that attempts at testing temperament traits such as persistance or endurance, by means of unpleasant tasks, (which, with adults at any rate may involve the possibility of electric shocks) lead to entirely artificial results. Carried to an extreme, in the effort to produce real life situations, this type of temperamental investigation seems to the present writer to result paradoxically in/
in isolated artificialities. She is convinced of the necessity of the 'total approach' which also must be essentially of a positive nature, as opposed to the negative aspect of the above mentioned type of test.

In the selection of tests, certain conditions had to be fulfilled. The first one was with regard to length. If there were to be any characteristic traits revealed, the test situation would have to last long enough so that the subject would become increasingly unaware of the test situation.

The question of difficulty next arose. With highly intelligent subjects most performance tests are not difficult enough to show reaction to difficulty. And even with the requisite degree of difficulty, the higher organization of the intelligent individual must be taken into account. The more complex and better integrated personality shows emotional tendencies in a more symbolized and less direct fashion than the less intelligent and the less well integrated. Also a high level of intelligence places the social adequacy of the individual so high that he can afford a greater degree of impairment without becoming socially ineffective. For this reason it is important to make the tests reach as difficult a level as possible.

Implicit in the difficulty level, is the matter of range. We wished to choose tests having a range of mental age from 8-16. There are very few existing performance tests which satisfy this criterion.

Finally it seemed important that the tests should be graduated.
graduated. They should be as free as possible from the element of chance. If the tests were graded in difficulty, an opportunity would arise for learning by experience. Also if the tests ranged sufficiently high in the intellectual scale, they should show up behaviour differences at various levels of ability in the same type of problem.

We should have mentioned also, that the time-factor is introduced in a lengthy test, which Peters\(^{(1)}\) considered so important in differentiating the psychopathic from the merely defective.

The writer considered that the order in which the tests were administered was of importance. The first test had to be sufficiently easy in its first stages to arouse confidence, and sufficiently attractive to stimulate interest. It was therefore decided to start off with a formboard. Two conditions had to be fulfilled here. First, the formboard had to be easily portable, for the tester could not very readily carry about too unwieldy a set of tests. This ruled out a formboard such as the Moorees. Second, the formboard had to fulfil the condition of length, and this disposed of the Oakley, which might perhaps have been the most obvious choice. The most suitable formboard for our purpose was therefore the Kent-Shakow, which in the writer's estimation, combines several virtues. It is lengthy, the material is attractive, it is well graded and embodies in the gradation a good principle not included in any other formboard known to the writer. Finally its compactness makes it easily portable.

I. Kent-Shakow formboard.

There are two models, the Clinical and the Industrial, the only difference being in size. The Industrial is larger, but both are essentially the same in design. No norms exist for the Industrial version, so the Clinical version plus already existing English norms was chosen. The test has been used in several reported American studies without any reference to publication of norms.

As can be seen from the diagrams, there are 6 sets of insets, for the same board. In the first set (2 S) the 5 shapes are to be filled in with wood divided into 2 and cut straight (hence 2 S.)

Similarly 2 D means that the wood is again divided into 2 parts, but diagonally cut (hence 2 D)

In the same way 3 S, 3 D, 4 S and 4 D mean increasing numbers of pieces cut straight first of all, and then diagonally. The wood is differently coloured at each stage. (The colours are indicated at each diagram).

In an article (1) by Shakow and Pazerian it was discovered that there are in addition 2 further sets of insets, 4 DD and 5 D respectively. Obviously this means an addition of more pieces, and if one had time to prolong the test and were not using it within a battery, it would be very useful. However, the English model not including those 2 stages has been used satisfactorily with adults, and the scoring method and norms are founded thereon.

The /

please send one copy of these to bring the first copy out
without giving a barrier. It would be very helpful. However,
the written form for identifying these & others has been used
satisfactorily with others, as the apparent method has shown me.

Tommy F. Evans
The actual formboard measures 11" by 8½". It consists of an uncut base of 3-ply wood. Screwed on to this base is the top piece of yellow pine, the 5 in cut shapes being of 6 mm. thickness. The insets themselves are 1 cm. thick.
Instructions. (1)

Tester and S. are seated at opposite sides of a table with the formboard before S. so that the horizontal figure is nearest him.

Tester places insets for 2 S on the table within easy reach of S. but so that the formboard is between the blocks and S. and says - "You see those holes ("shapes" with older subjects) - what you do is to fill them up with the pieces of wood. They all fit in somewhere, and none will be left over". (No mention was made of speed, for it was part of S's intention to watch individual rates.)

Pieces are shuffled and placed before S. at random. No two adjacent pieces should be placed together in any task. (The writer actually put out the pieces in the same order always, so that a chance piece would not occur to help one subject more than another.)

The tasks are presented in the following order: - 2 S, 2 D, 3 S, 3 D, 4 S, 4 D, until S. fails either by giving up or by exceeding the time limits. Failure in any diagonal task indicates S's limits in the diagonal series, but he should still be permitted to try another task of the straight series, if there are any such left. An S. who loses credit on 3 D by exceeding the time limit may yet be able to score on 4 S.

Although no credit is given for a performance that passes the time limit, S. should be permitted, if there are any indications that he can, to finish the task unaided. On no account should he be permitted to suspect that there are any time limits.

<table>
<thead>
<tr>
<th>Time Interval</th>
<th>2S</th>
<th>2D</th>
<th>3S</th>
<th>3D</th>
<th>4S</th>
<th>4D</th>
</tr>
</thead>
<tbody>
<tr>
<td>To 15&quot;</td>
<td>6</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 - 30&quot;</td>
<td>5</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31 - 45&quot;</td>
<td>4</td>
<td>6</td>
<td>10</td>
<td>16</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>46 - 60&quot;</td>
<td>3</td>
<td>5</td>
<td>9</td>
<td>15</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1:01 - 1:30&quot;</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>14</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>1:31 - 2:00</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>13</td>
<td>9</td>
<td>19</td>
</tr>
<tr>
<td>2:01 - 2:30</td>
<td>6</td>
<td>12</td>
<td>8</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:31 - 3:00</td>
<td>5</td>
<td>11</td>
<td>7</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3:01 - 4:00</td>
<td>10</td>
<td>6</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4:01 - 5:00</td>
<td>9</td>
<td></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5:01 - 6:00</td>
<td></td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6:01 - 7:00</td>
<td></td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NORMS.

- 6:31 - 7:00, 13+
- 8:00 - 9:00, 52-55
- 9:00 up

May 22, 1951.
II. Cube Construction Test.

This test was given next because, coming after the rather lengthy formboard, it seemed short, and provided a kind of break between Kent-Shakow and P.C.C. This test is described by Gaw\(^1\) and by Collins and Drever\(^2\).

This test is so well known that it does not require to be described in such detail as the much less known Kent-Shakow formboard.

It is in three parts. Part one comprises the making by the subject of the middle section of a three-inch cube painted grey\(^3\) outside. This section is shown to the subject, being painted grey round the edges, and he is given the necessary nine one-inch cubes suitably painted to copy the model. In part two the section is painted grey on one side and round the edges (being the top section of the 3-inch cube already referred to), and again the subject is given suitably painted one-inch cubes to reproduce the model. In part three the model is a two-inch cube with no paint on the outside surface. The subject is given eight one-inch cubes all painted grey on three surfaces and is required to produce a two-inch cube with no paint showing on the outside surfaces at all.

The method of administration was that given by Drever and Collins, but the method of scoring used was that devised by Alexander\(^4\). This was done to increase the discrimination at the/
the higher intelligence levels. The criteria used were number of cubes correctly placed and time required in doing the test. Also the time score was made to depend on the fraction of success.

Instructions.

Show Model I to the subject, explaining that it is painted on the sides, but not on the top or the bottom. Place the model on the table and demonstrate how "one just the same" can be made with the blocks provided. On completion of the model lift it to show that it is just the same. Scatter the blocks and tell the subject to make one just the same. Note the time in seconds.

Show Model II to the subject, explaining that it is painted on the sides and top but not on the bottom. Place it before the subject, scatter the blocks and tell him to make one "just the same." Note time in seconds.

Show Model III to the subject, explaining that it is not painted at all. Show the subject that all the blocks are the same, three sides being painted, and three unpainted. Then say "Now I want you to make one just the same as this (point) with these blocks. Go ahead."

Method of Scoring.

A subject may finish to his own satisfaction and be only partially successful. To overcome this the time score is in each case multiplied by the "power" score and divided by the possible power score.

Part I./
Part I  Power Score.  1 for each cube correct 9
    Time Score. Take, from 10, 1 for each 3\(\frac{1}{4}\) min. needed 10
    Total 19

Part II  Power Score.  2 for each cube correct 18
    Time Score. Take, from 20, 1 for each 3\(\frac{1}{4}\) min. needed 20
    Total 38

Part III  Power Score.  3 for each cube correct 24
    Time Score. Take, from 20, 1 for each 3\(\frac{1}{4}\) min. needed 20
    Total 44

Time Score found from the above formula should be multiplied by the fraction. Actual power score/Possible power score. For example, take part 3 and the case where there are four blocks correct (a common case). The power score is then 12 and the time score is obtained from the time taken; say this is 68 sec., the score is 16 multiplied by 12/24, i.e. 8. The score then for part 3 is 12 + 8, i.e. 20

III. P.C.C. test.

Instructions and Method of Scoring having been already described in Ch. III.

IV. Healy Picture Completion II.

This test was given at this point to vary the interest of the battery, for the first three tests had been essentially concerned/
TABLE OF TIME SCORES FOR THE CUBE CONSTRUCTION TEST.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>19</td>
<td>18</td>
<td>17</td>
<td>16</td>
<td>15</td>
<td>14</td>
<td>13</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>19</td>
<td>18</td>
<td>17</td>
<td>16</td>
<td>15</td>
<td>14</td>
<td>13</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
concerned with manipulating bits of wood. The more imaginative picture completion seemed just right from the point of view of variety and slight relaxation. (1)

This was administered, as by Gaw.

Instructions.
"Here is a picture - it begins here (pointing to demonstration picture) where the boy is getting dressed. It shows the same boy doing one thing after another during the same day." (Point along the rows to indicate course of pictures.) "You see in each picture that a piece is missing. Here (pointing to them) are many small pieces, far more than you need, and they all fit the spaces. What you have to do is to pick out the one you think suits the picture best. For instance, what is needed here? (pointing to demonstration picture). Yes, a boot." If an incorrect answer is given, say, "No, he is dressing, and he is bending down for his other boot. Now, which is the best piece for us to put in?" If a shoe is selected instead of a boot explain: "This one would not be right as he must have a boot to go with his other boot."

Go ahead with the others. And remember to choose the best piece for each space.

Scoring:/

Scoring:

<table>
<thead>
<tr>
<th>No. of Picture</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4(8) 1(12) 2(30) 1(31) 3(45) 1(56)</td>
</tr>
<tr>
<td>2</td>
<td>5(19) 5(13) 2(19) 5(27) 5(47) 1(49)</td>
</tr>
<tr>
<td>3</td>
<td>4(13) 1(15) 1(35) 1(36) 4(37) 1(44) 6(46) 1(58) 2(40) 1(52) 6(52)</td>
</tr>
<tr>
<td>4</td>
<td>3(13) 3(7) 2(13) 1(16) 3(11) 13(23) 1(30) 8(31) 3(42) 2(45) 2(57) 2(53) 1(58)</td>
</tr>
<tr>
<td>5</td>
<td>1(13) 3(13) 1(21) 2(32) 1(41) 7(32) 2(41) 2(42) 1(44) 1(47) 1(49)</td>
</tr>
<tr>
<td>6</td>
<td>1(14) 4(34) 0(25) 0(36) 1(44)</td>
</tr>
<tr>
<td>7</td>
<td>4(34) 6(17) 7(33)</td>
</tr>
<tr>
<td>8</td>
<td>6(13) 6(17) 7(33)</td>
</tr>
<tr>
<td>9</td>
<td>1(39) 1(50) 1(57) 2(59)</td>
</tr>
<tr>
<td>10</td>
<td>7(13) 6(44) 1(42) 1(13) 6(22) 6(24) 1(48) 2(47) 2(67)</td>
</tr>
</tbody>
</table>

The figure in brackets refers to the number of the insert placed, the accompanying figure is the score credited in each case.
V. **Pass-along** (Alexander).

Alexander's Pass-along test comprises 9 subjects of increasing difficulty. In each case the subject is presented with a box in which are placed a number of wooden blocks. The task is to transfer one of these blocks from one side of the box to the other (colours are used to make the task clear and a diagram of the end position is always before the subject.) In the solution the subject must not lift the blocks out of the box but must move them in the limited space that remains in the box. There is no great manual difficulty involved. The order of difficulty increases from the first box, on which the average 5-year old could probably achieve success, to the ninth box on which only superior adults have shown success. One of the greatest assets in using this test is the facility with which interest is maintained in its administration.

The method of scoring is the second one devised by Alexander. Time is taken more carefully into account and the total score depends both on success and on time. The possible score is increased to 102 instead of 45.

**Instructions.**

Along with the above material, there is a set of diagrams showing the end positions for each of the sub-tests. To get the initial position we simply consider the painted ends of the box as being reversed. Alternatively lift the red block to the blue end pushing the blue blocks to the red end. There is one exception to this rule, in the case of the third subtest. The end position for/

(1) B.J.P. vol XXIII Pt.I 1932.
for subtest, or box 3 is the same as for box 2. The initial position for box 3 is obtained by placing the two red blocks at the blue end with a blue block between them, and putting the remaining two blue blocks at the red end.

To gain the test, we place diagram 1 before the subject and box 1 in its initial position. Point out to the subject that in the picture, the red block is at the red end and the blue blocks at the blue end. Then say "Now, make this the same as the picture". Show S. that the blocks can be moved, and demonstrate by moving one of them. Warn the subject that they must not be lifted out of the box. Note the time in seconds. The purpose of this first box is to introduce the subject to the test material and to make sure that he knows what is required of him. If necessary allow several chances with the first box. In each case place diagram and box before the subject and say "Go ahead and move the red blocks to the red end so that it will be like the picture". As a rule it will be found unnecessary to repeat instructions. In boxes 8 and 9 draw the attention of S. to the position of the small blue blocks, explaining that they must be as in the picture at the end. Apart from this no instructions are needed. It is necessary to avoid exposing two diagrams simultaneously, as this may help the subject, since the tests are arranged so that each leads to the next one.
TABLE OF TIME SCORES FOR THE PASSALONG TEST.

Time taken in Sec.

<table>
<thead>
<tr>
<th>No.</th>
<th>15</th>
<th>30</th>
<th>45</th>
<th>60</th>
<th>75</th>
<th>90</th>
<th>105</th>
<th>120</th>
<th>135</th>
<th>150</th>
<th>165</th>
<th>180</th>
<th>195</th>
<th>210</th>
<th>225</th>
<th>240</th>
<th>255</th>
<th>270</th>
<th>285</th>
<th>300</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
</tr>
</tbody>
</table>
VI. Kohs Block Design Test.

Ten patterns of the test, as selected in the Drever & Collins scale, were used. Again Alexander's method of scoring was used, in order to increase discrimination at the upper levels. This method of scoring takes into account success and time. The test makes use of 16 one-inch cubes which are painted on the sides. By means of these cubes various patterns in colour can be produced. Each cube has one side yellow, one red, one blue, and one white, one blue and yellow diagonally divided, the other red and white diagonally divided. The tests use at first only four blocks, then nine blocks, and finally sixteen. A diagram showing the required pattern is placed before the subject for his constant guidance during each test.

Instructions.

Take four of the blocks and place them white side up before the subject. Then turn the blocks (or get the subject to turn them), showing in succession the blue, the yellow, the red, the red and white, and the blue and yellow sides. This is to show the subject the colour or colours on each side, and also to make clear to him that all the blocks are the same.

(1) Place design 1 and place before S. Show him how to make design 1 from the four blocks. Scatter the blocks and turn them all over. Ask S. to put them together again to form the design. If he fails to respond or does it wrongly show again, and allow a second chance.

(2) Place design 2 before S. and scatter blocks so that only one shows/
shows the right side up. Allow 90 sec. for design. If unsuccessful, demonstrate, and go on to design 3, without allowing a second trial.

(3) Proceed as before. When two successive failures occur, stop the test. If failure occurs on any one, demonstrate and go on to the next.

(4) and (5) are administered as (3).

After (5) add five more blocks, showing S. that they are exactly the same as the others. Scatter so that only two show the right side up and proceed to (6) and (7). Add seven more blocks, showing as before that they are all the same. Proceed to designs (8), (9) and (10).

As has already been stated, the writer administered the test battery strictly in the order described. She found that the tests were thus graded and varied as much as possible, relative to the material. The battery was given complete to each individual and this variety, along with short pauses as each new test was taken out, interspersed with a few pleasant remarks, appeared to have avoided undue fatigue on the part of the subject.
B. Group Tests used in the Investigation.

I. N.I.I.P. Space Perception Test.

This test was devised by Slater as a measure of k-factor. It is divided into four parts, and in each part the essential problem is to choose out of several presented shapes the one, which matches the specimen one.

Under the instructions for each separate part we have given a specimen of the appropriate section.

Test I. Instructions.

When name and other information have been filled in, say "Turn over to page one. I will read the instructions aloud. While I do this I want you to follow them carefully on your own books." Read the instructions through. Allow half a minute (or more if necessary) for the two examples. Then say, "The answer to the first problem is 4. Number 2 won't do, because the notch is up at the top of the circle. Number 3 won't do because it is down at the bottom. Number 1 won't do because although the notch is on the proper side of the circle it is too high up. Number 4 is the only one which looks like the one at the beginning of the row turned over sideways. The answer to the second problem is 3, because it is the only one of the four numbered shapes which has its legs pointing in the proper direction. The problems on the next three pages are just like these. You will have five minutes for the whole of the three pages. Turn over straight away to the next page and begin. Don't forget there are three pages to be done."

After five minutes say, "Stop."
Test 2.

"Now turn over to the next test - Test 2. I will read through the instructions again." Read the instructions and allow time for the working of the two examples. Then say, "The answer to the first problem is 1, because Number 1 shape is the only one which is like the one at the beginning of the row turned over sideways. The others are the wrong shape and size. The answer to the second problem is 3, because Number 3 shape is the only one which is like the one at the beginning of the row turned over sideways. That is a harder one and some of you may have got it wrong, but if you look at it carefully you will see that Number 3 is the only one which is the right shape and size. The problems on the next three pages are just like these. You will have five minutes for the whole of the next three pages. Turn over straight away and begin. Don't forget that there are three pages to be done." After five minutes say, "Stop."
Test 3.

"Now turn over to the next test - Test 3. I will read through the instructions again." Read the instructions and allow time for the working of the two examples. Then say, "The answer to the first problem is 1, because Number 1 shape is the only one of the four numbered shapes which has long enough sides and a blunt enough point. The answer to the second problem is 3, because it is the only one which has the same shape and size as the one at the beginning of the line. The problems on the next three pages are like these ones. You will have five minutes for the whole of the three pages. Turn over straight away and begin. Don't forget that there are three pages to be done."

After five minutes, say, "Stop."
Test 4.

"Now turn over to the last test - Test 4. I will read through the instructions again." Read through the instructions and allow time for the working of the two examples. Then say, "The answer to the first problem is 1. Number 2 shape is like the one at the beginning of the row when it has been turned half way round; and Number 3 shape is like the one at the beginning when it has been turned three quarters of the way round. Number 4 shape is like the one at the beginning before it has been turned round at all. But Number 1 is the only shape which is like the one at the beginning of the row when it has been turned over as well as round. The answer to the second problem is 3. This is the only shape which is like the one at the beginning of/
of the row when it has been turned over as well as round. The problems on the next three pages are like these ones. You will have five minutes for the whole of the next three pages. Turn over straight away and begin. Don't forget that there are three pages to be done." After five minutes, say, "Stop now and close your books."

**Scoring 1 for each answer correct. Total 60.**

---

**II. Progressive Matrices (1938)**

This is a non-verbal series by Raven intended to measure intelligence. The tests are held to be suitable for almost any testee, irrespective of age, linguistic ability, or physical defects.

The printed form of the series is designed as an individual or group test and is suitable for children above 6 years of age and for adults.

Each problem consists of a design or "Matrix" from which part has been removed. A testee has to examine the matrix and decide which of the pieces given below is the right one to complete it. A testee (in the group test) indicates his choice by writing down on a scoring form the number of the piece he selects.

(1) "A new series of Perceptual Tests: Preliminary Communication."


Standardisation of Progressive Matrices."


* This test has now been withdrawn from general circulation. It will therefore be described very briefly.
selects.

There are five sets of problems. Each set develops a different theme. The initial tests in each set are easy so as to be self-evident and these are followed by tests of increasing difficulty; the order in which they are presented provides the necessary training. If a testee does not at first grasp the nature of the task the explanation may be altered.

From the results the psychologist can classify a testee according to where his score falls on a percentile scale. The reliability of a testee's score is indicated by his scores on the five component sets. The time taken and the nature of the testee's errors are held to be indicative of temperamental and emotional traits.

Instructions.

"Open your books to the first page. It is like this." (Demonstrates.) "At the top it says Set A and you have a column A here, on your scoring form. This is A 1. You see what it is. The upper part is a pattern with a bit left out. Each of these bits below (points to each in turn) is the right shape to fit the space but they do not all complete the pattern. Number 1 (he points to the bit and then to the pattern) is quite the wrong pattern. Numbers 2 and 3 are wrong - they fit the space but they are not the right pattern. What about number 6? It is the right pattern (illustrates that pattern is the same as the pattern above) but it does not go all over. Put your finger/
finger on the one that is quite right." Further explanation is
given, if necessary. "Yes, number 4 is the right one. So the
answer to A 1 is 4 - write 4 here, against number 1 in Column
A on your scoring form."

"On every page in your book there is a pattern with a bit
left out. All you have to do is to choose each time which of
the bits is the right one to complete the pattern. When you
have found the right bit you write the number of it down on
your scoring form against the number of the pattern. They are
simple at the beginning and get harder as you go on. There is
no catch. If you pay attention to the way the easy ones go,
you will find the later ones less difficult. Try each in turn,
right to the end of the book. Go your own pace. Do not miss
any out. Do not turn back. See how many you can get right.
You can have as much time as you like. Turn over and do the
next one."

Scoring.

1 for each correct solution. Total 60.

III. Verbal Group Tests.

(1) Trabue Completion-Test Language Scale. Form L.

(2) Trabue Completion-Test Language Scale. Form M.

(3) Vocabulary Test.

The Trabue Sentence Completion tests consist essentially of
sentences from which words are missed out. S. fills in what he
thinks are the missing words. Form L and M each contain 8
sentences, for example:
The ...... seems ...... and dreary ...... a discouraged ...... .
Only/
Only one word is to be written in each blank. The Time Limit for both L and M is five minutes.

Instructions for Sentence Completion L and M.

"This sheet contains some incomplete sentences - sentences in which some of the words were left out. There is a blank in each place where a word should be written. You are to write one word on each blank, in each case writing the word which makes the most sensible statement. You will have just 5 mins. in which to write your name and age at the top of the page and write the words that are missing. Remember that you are to write just one word on each blank."

Scoring.
Score of 2 for each perfectly completed sentence. Errors in spelling, capitalization or punctuation not allowed to lower score.
Score of 1 for each sentence with only a slight imperfection - poorly chosen word or common grammatical error which makes sentence less than perfect yet leaves it with reasonably good sense.
Score of 0 for a sentence with sense or construction badly distorted. Many examples grouped under the 3 different scores are given by Trabue (1).

Vocabulary Test.
A word meaning the same, or nearly the same, beginning with C must be found for 25 different words which increase in difficulty.
For/

(1) Key for Completion-Test Language Scales. Teachers College, Columbia University, New York, 1922.
For example, a word must be found beginning with C which means the same as "bodily". (Another example was "waterfall").

Instructions:
"Here are some words. You are to think of other words which mean the same or nearly the same, and which begin with "C". Every word you write must begin with "C." Write your words on the lines opposite the words to which they belong, like this:

You will be allowed 10 minutes.

Greed  -  Covetousness
A Prisoner  -  Convict

Score.
2 for each correct. Total 50.

Order of Administration of Group Tests.

The tests were actually given in the following order: Matrices, Space Perception, Verbal battery. There was a reason for this order. As we have already mentioned, as little attention as possible was paid to time throughout the performance tests. Scores were calculated by time limits, but only if S. was grossly over-running the limit was S. stopped or made to go on to the next step. It seemed to the writer that to introduce a speed factor into the subjects' awareness, was going to perhaps change, or at least to influence the motivation. What she wished to observe was the native rate of the Subject, if that were possible.

The test, Progressive Matrices, was thus used with its original instructions of limitless time, and not with its later conditions of timing. It was therefore given first of the group/
group tests, being in line with the other lack of time limits.

Both Space Perception and the Verbal Battery, however, had to be finished strictly to time. For one thing it would have taken far too long to wait for each individual to finish, since Space Perception itself takes about 25-30 minutes to administer, and the verbal battery a similar time. Also it was impossible to record each individual's time for him as he finished, for the sections of the S.P test were too short. This was possible with Matrices, however. Each subject held up his hand as he finished and the tester recorded his time. The two tests involving time limits were thus kept to the end so that any incitement to speed would not be carried over to other tests where it was not wanted.

Each of the Group tests was given on a separate day.
Ch. VI. **GENERAL RESULTS WITH COMPLETE TEST BATTERY.**

The Group tested.

The writer decided that it was essential to do all the testing herself, for as it has been more than once mentioned, she wished to rely to a great extent on direct observation during the tests. Since the administration of a battery of performance tests individually is very time consuming, the number of cases had therefore to be limited.

The investigator decided to select a group including in age boys and girls of 13 years who were not yet 14. Since the primary purpose was to observe temperament reactions, it was considered important that intelligence should be controlled. The group had therefore to be as homogeneous as possible in intelligence. It was realised that this homogeneity would lower the correlation coefficients between the tests, but the importance of keeping constant as many factors as possible in the temperamental situation seemed to outweigh this disadvantage.

A group of 50 children from Leith Academy Secondary School (25 boys and 25 girls) was therefore chosen, whose ages fell within one year-group, and who were all in the A section of their class, their I.Q.'s ranging thus from superior to very superior. The I.Q.'s had been recorded at the Qualifying Examination period (the English 11+ stage) on the result of a Moray House Test. The mean I.Q. was 118, and the range was from 110 to 135. Even this was a considerable range, but it was the best that could be done under the circumstances.

It was very unfortunate that after the 50 subjects had done all the tests, except the final verbal battery, a boy and a girl who had just reached the age of 14 at that time decided to leave school. Although both missed only the verbal tests out of the entire /
entire battery, it was decided to discount all their scores from the results. The results are therefore based on 48 cases, (24 boys and 24 girls).

**Test Distributions.**

On the following pages will be found the distributions for the 9 tests. Although there were 3 separate verbal tests, they are combined into one for statistical purposes. The tests are Kent-Shakow formboard, Cube Construction, P.C.C., Healy II, Passalong, Kohs blocks; Progressive Matrices, N.I.I.P. Space Perception, Verbal battery. In all cases the raw scores only were used. Neither I.Q. nor Mental Age was necessary for the calculations. In any case norms for P.C.C. have not yet been established, and norms for N.I.I.P. Space Perception were not available.
Cube Construction

N = 48

<table>
<thead>
<tr>
<th>Group</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-10</td>
<td>1</td>
</tr>
<tr>
<td>11-15</td>
<td>1</td>
</tr>
<tr>
<td>16-20</td>
<td>3</td>
</tr>
<tr>
<td>21-25</td>
<td>4</td>
</tr>
<tr>
<td>26-30</td>
<td>5</td>
</tr>
<tr>
<td>31-35</td>
<td>4</td>
</tr>
<tr>
<td>36-40</td>
<td>3</td>
</tr>
<tr>
<td>41-45</td>
<td>2</td>
</tr>
<tr>
<td>46-50</td>
<td>1</td>
</tr>
</tbody>
</table>

P.C.C

N = 48

<table>
<thead>
<tr>
<th>Group</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>1</td>
</tr>
<tr>
<td>6-10</td>
<td>3</td>
</tr>
<tr>
<td>11-15</td>
<td>4</td>
</tr>
<tr>
<td>16-20</td>
<td>6</td>
</tr>
<tr>
<td>21-25</td>
<td>4</td>
</tr>
<tr>
<td>26-30</td>
<td>5</td>
</tr>
<tr>
<td>31-35</td>
<td>4</td>
</tr>
<tr>
<td>36-40</td>
<td>5</td>
</tr>
<tr>
<td>41-45</td>
<td>3</td>
</tr>
<tr>
<td>46-50</td>
<td>1</td>
</tr>
</tbody>
</table>
HEALY PICTURE COMPLETION I

N = 48

PASSALONG

N = 48
Means and Standard Deviations of Raw Scores.

<table>
<thead>
<tr>
<th>Test</th>
<th>N</th>
<th>Possible Total</th>
<th>Mean (M)</th>
<th>Standard Error (σ_m)</th>
<th>Standard Deviation (SD)</th>
<th>Standard Error (σ SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-S fm. bd.</td>
<td>72</td>
<td>43.94</td>
<td>1.56</td>
<td>10.37</td>
<td>1.12</td>
<td></td>
</tr>
<tr>
<td>C.C.</td>
<td>101</td>
<td>33.30</td>
<td>1.09</td>
<td>7.55</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>P.C.C.</td>
<td>50</td>
<td>26.72</td>
<td>1.70</td>
<td>11.80</td>
<td>1.21</td>
<td></td>
</tr>
<tr>
<td>Healy</td>
<td>97</td>
<td>69.06</td>
<td>1.97</td>
<td>13.69</td>
<td>1.41</td>
<td></td>
</tr>
<tr>
<td>Passalong</td>
<td>102</td>
<td>55.56</td>
<td>1.52</td>
<td>10.56</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>Kohs</td>
<td>93</td>
<td>58.24</td>
<td>2.69</td>
<td>18.64</td>
<td>1.92</td>
<td></td>
</tr>
<tr>
<td>Matrices</td>
<td>60</td>
<td>49.04</td>
<td>.73</td>
<td>5.05</td>
<td>.52</td>
<td></td>
</tr>
<tr>
<td>Space Perc.</td>
<td>80</td>
<td>40.31</td>
<td>1.17</td>
<td>8.13</td>
<td>.83</td>
<td></td>
</tr>
<tr>
<td>3 Verbal</td>
<td>82</td>
<td>33.46</td>
<td>1.16</td>
<td>8.04</td>
<td>.82</td>
<td></td>
</tr>
</tbody>
</table>

As can be seen from the distributions, the scores tend to be rather bunched together, as we might have expected from the narrow nature of our sample. Within their own range, however, the distributions approximate to the normal.

Inter-correlations of 9 tests.

If we look at the correlation matrix, we see that all the performance tests, except Passalong and Healy Picture Completion, are significantly correlated at the 1% level. From Fisher’s (1) tables we see that for 48 cases, an r would have to be .279 to be significant at the 5% level, and .361 to be significant at the 1% level.

Passalong correlates significantly with only one test (Cube Construction) and even this might possibly be attributed to chance. /

## INTER-CORRELATIONS OF 9 TESTS.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>K.S.</td>
<td>I</td>
<td>II</td>
<td>III</td>
<td>IV</td>
<td>V</td>
<td>VI</td>
<td>VII</td>
<td>VIII</td>
<td>IX</td>
</tr>
<tr>
<td>I</td>
<td>-</td>
<td>.559</td>
<td>.479</td>
<td>.199</td>
<td>.119</td>
<td>.626</td>
<td>.536</td>
<td>.601</td>
<td>.027</td>
</tr>
<tr>
<td>C.C.</td>
<td>II</td>
<td>-</td>
<td>.535</td>
<td>.057</td>
<td>.499</td>
<td>.586</td>
<td>.553</td>
<td>.499</td>
<td>.000</td>
</tr>
<tr>
<td>P.C.C.</td>
<td>III</td>
<td>.479</td>
<td>-</td>
<td>.214</td>
<td>.258</td>
<td>.569</td>
<td>.493</td>
<td>.535</td>
<td>-.031</td>
</tr>
<tr>
<td>Healy II</td>
<td>IV</td>
<td>.109</td>
<td>.057</td>
<td>-</td>
<td>-.160</td>
<td>.215</td>
<td>.130</td>
<td>.227</td>
<td>.059</td>
</tr>
<tr>
<td>Pass.</td>
<td>V</td>
<td>.110</td>
<td>.409</td>
<td>.258</td>
<td>-.160</td>
<td>-</td>
<td>.136</td>
<td>.105</td>
<td>.085</td>
</tr>
<tr>
<td>Kohs</td>
<td>VI</td>
<td>.526</td>
<td>.598</td>
<td>.659</td>
<td>.215</td>
<td>.136</td>
<td>-</td>
<td>.582</td>
<td>.633</td>
</tr>
<tr>
<td>Mat.</td>
<td>VII</td>
<td>.536</td>
<td>.553</td>
<td>.493</td>
<td>.130</td>
<td>.103</td>
<td>.582</td>
<td>-</td>
<td>.100</td>
</tr>
<tr>
<td>S.P.</td>
<td>VIII</td>
<td>.601</td>
<td>.499</td>
<td>.535</td>
<td>.227</td>
<td>.085</td>
<td>.633</td>
<td>.592</td>
<td>-</td>
</tr>
<tr>
<td>V.</td>
<td>IX</td>
<td>.027</td>
<td>.000</td>
<td>-.031</td>
<td>.039</td>
<td>-.092</td>
<td>.049</td>
<td>.100</td>
<td>-.030</td>
</tr>
</tbody>
</table>

\[ N = 48 \]
\[ 1\% \text{ significance level of } r \text{ is } .361 \]
\[ 5\% \text{ " " " " " " } .279 \]

*Fisher's Tables.*

---

I. Kent-Shakow form board.
II. Cube Construction.
III. P.C.C.
IV. Healy Picture Completion II.
V. Pass-along
VI. Kohs blocks
VII. Progressive Matrices
VIII. M.I.I.P. Space Perception Test.
IX. Verbal battery
   (2 sentence completion)
   (1 Vocabulary,
chance. It was surprising to find that Passalong did not fall in with the other performance tests, but the writer thinks this may be explained on the grounds of its inapplicability to superior children of 13. Most of the pupils reached step 7 but could not complete step 8. This point (between 7 and 8) seems to be the really crucial point of the whole test. Two children managed to do step 8, but this appeared to be due to chance. Neither of them seemed to have grasped any principle which would have enabled them to reach the final step, and of course step 9 is actually done on the same principle as step 8. The differences in score up till step 7 were thus due to differences in time and not in achievement.

The lack of correlation of Healy with the other performance tests was, however, no great surprise to the writer. She is of the opinion that a picture completion test simply is not a performance test; at any rate not in the sense in which a formboard, or Kohs blocks are performance tests. We shall return to this point later.

The Kent-Snake formboard, P.C.C., Cube Construction and Kohs blocks are all significantly inter-correlated. In fact the uniformity of those rs is quite striking, and they are fairly high, also, considering the homogeneity of the sample. The highest rs is that between Kohs and P.C.C. (.659). Amongst the performance test rs Kohs showed highest, with P.C.C. next.

Matrices and Space Perception likewise show significant rs with the four performance tests. The same uniformity is also in evidence.

The rs of the Verbal tests were rather surprising. They too show uniformity in that they are all zero, except for the r .100! They were chosen with the idea in mind of having as pure /
pure an estimate of "v" as possible within the battery in order to see how verbal tests would compare with performance. It was anticipated, however, that there would be some degree of positive correlation, since we might assume a measure of g common to both performance and verbal tests. We correlated the verbal scores with school achievement in English and found \( r = .314 \) (significant at 5% level but not at 1%). This did not really tell us very much though. It did seem to the writer that the sentence completion type of test, having to be done within a very strict time limit, would measure fairly satisfactorily the kind of verbal fluency which we associate with "v".

She was not so sure that the same could be said for the vocabulary test (which was the third of the verbal battery), for vocabulary has been held by some to correlate highly with intelligence; Terman, for instance, found that it correlated higher with the whole Revision scale than any other item. It did seem proper, however, to include a vocabulary test in our short verbal battery. But then, perhaps the "purity" of "v" functioning along with "g" to give us verbal ability is a fiction. Possibly "v" includes within itself several verbal factors.

If the verbal tests had showed higher rs we might have used a Holzinger Bi-factor analysis with the data.

We ought to add, too, that on the whole the verbal scores tended to be poor. Compare them with scores on the Matrices, for instance, where the mean score is 49, out of a possible total of 60. Linguistic tests such as the Binet have always been biased by the individual's linguistic background. We believe that the verbal scores with the present sample were probably similarly affected. The home background tends to be medium to superior working-class. The spelling, also, in the vocabulary test, was below the standard one might have expected from the intelligence/
intelligence status of the group.

**Factorial Analysis. (Centroid).**

We eliminated the tests which showed non-significant rs and were thus left with the remaining 6 tests -

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>.626</td>
<td>.559</td>
<td>.470</td>
<td>.626</td>
<td>.601</td>
<td>.536</td>
</tr>
<tr>
<td>II</td>
<td>.559</td>
<td>.596</td>
<td>.535</td>
<td>.596</td>
<td>.499</td>
<td>.553</td>
</tr>
<tr>
<td>III</td>
<td>.470</td>
<td>.535</td>
<td>.659</td>
<td>.659</td>
<td>.535</td>
<td>.493</td>
</tr>
<tr>
<td>IV</td>
<td>.526</td>
<td>.596</td>
<td>.659</td>
<td>.659</td>
<td>.633</td>
<td>.582</td>
</tr>
<tr>
<td>V</td>
<td>.601</td>
<td>.499</td>
<td>.535</td>
<td>.633</td>
<td>.633</td>
<td>.592</td>
</tr>
<tr>
<td>VI</td>
<td>.536</td>
<td>.553</td>
<td>.493</td>
<td>.582</td>
<td>.592</td>
<td></td>
</tr>
</tbody>
</table>

If we insert in the diagonal cells of the above matrix the highest rs in the corresponding row and column, and then extract Thurstone's first centroid factor, we find the following loadings of the 6 tests in this factor -

<table>
<thead>
<tr>
<th>Tests</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loading</td>
<td>.751</td>
<td>.734</td>
<td>.737</td>
<td>.825</td>
<td>.763</td>
<td>.736</td>
</tr>
</tbody>
</table>

Since we found the first factor loadings to be so large, we decided that rather than attempt to take out another factor, we would instead see if the matrix of correlations could not be explained in terms of one general factor. This does not mean that only one factor is necessarily present. It merely means that it is sufficient to explain the data in terms of one factor.
In order to test this hypothesis, we make use of the Method of Maximum Likelihood, devised by D. N. Lawley (Education Dept., Edinburgh University) to estimate the significance of factor loadings.

Applying the Maximum Likelihood method of factor estimation we obtain the following values for the factor loadings:

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loading</td>
<td>.743</td>
<td>.718</td>
<td>.720</td>
<td>.844</td>
<td>.762</td>
<td>.723</td>
</tr>
<tr>
<td>Community</td>
<td>.552</td>
<td>.516</td>
<td>.513</td>
<td>.712</td>
<td>.531</td>
<td>.523</td>
</tr>
<tr>
<td>Specific Variance</td>
<td>.448</td>
<td>.434</td>
<td>.482</td>
<td>.288</td>
<td>.419</td>
<td>.473</td>
</tr>
</tbody>
</table>

(1 - \(\ell^2\))

The residuals left after extracting the effect of the general factor from the correlations are then given in the table below:

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>((.448))</td>
<td>.026</td>
<td>-.065</td>
<td>-.001</td>
<td>.035</td>
<td>-.001</td>
<td>I</td>
</tr>
<tr>
<td>((.434))</td>
<td>.018</td>
<td>-.010</td>
<td>-.043</td>
<td>.034</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>((.432))</td>
<td>.051</td>
<td>-.014</td>
<td>-.028</td>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>((.283))</td>
<td>-.010</td>
<td>-.023</td>
<td>IV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>((.419))</td>
<td>.041</td>
<td>V</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(.477)</td>
<td>VI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(We have inserted in the diagonal cells, the specific variances, i.e. \(1 - \(\ell^2\)\)).

It is now possible to calculate a measure of discrepancy between hypothesis and observation. This is given by

\[
W = 48 \left\{ \frac{(.026)^2}{.448 \times .434} + \frac{(-.065)^2}{.448 \times .432} + \cdots \text{(15 terms)} \right\}
\]

\[
= 48 \times .0858
\]

\[
= 4.12
\]
To find whether this value of \( w \) is significant or not we use the \( \chi^2 \)-table with \( p \) degrees of freedom, where

\[
p = \frac{1}{2} \left\{ (6-1)^2 - 6 - 1 \right\} = 9
\]

The 5% point for \( \chi^2 \) with 9 d.f. is 16.92, so the value of 4.12 obtained for \( w \) is not significant. There are, in fact, 90 chances in 100 of getting a value for \( w \) greater than 4.12.

The original hypothesis is therefore borne out by the above calculation, so we may assume the existence of only one general factor.

Clearly this method of factor estimation is of considerable importance for factorial analysis. It means that many of the reported researches using the centroid method have assumed the existence of significant factors which, if this method were applied, would prove to be not significant. The existence of several factors is usually assumed after using Thurstone's method of rotating the axis for each factor, in order to endow the factors with psychological significance. Thurstone's rotation is essentially an arbitrary procedure, but in the light of the results of the preceding method, it seems even more so.

It would be interesting to apply the method to Alexander's test results. Let us recall just four of the tests he used, in order to compare the factor loadings with our own.

<table>
<thead>
<tr>
<th>Tests</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>( h^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passalong</td>
<td>.608</td>
<td>-.336</td>
<td>-.15</td>
<td>.426</td>
</tr>
<tr>
<td>Kohs</td>
<td>.758</td>
<td>-.393</td>
<td>-.098</td>
<td>.698</td>
</tr>
<tr>
<td>Cube Construction</td>
<td>.670</td>
<td>-.414</td>
<td>-.098</td>
<td>.630</td>
</tr>
<tr>
<td>Formboards</td>
<td>.600</td>
<td>-.453</td>
<td>-.275</td>
<td>.641</td>
</tr>
</tbody>
</table>

\( \text{Factor Loadings.} \)

\( (1) \) op. cit. pp. 90.

This method is intended for use with large samples, but it is quite applicable here since \( w \) is nowhere near being significant.
Alexander attached no psychological significance to these factors as they stand, since he meant to rotate the axis so that each factor would be "defined according to psychological concepts already accepted".

Although Alexander gave a highly interesting (and to the present writer, sympathetic) account of the psychological significance of his factors, once they had been rotated, it is yet hard to accept his conclusions as proved. His conclusions remain hypothetical, for to rotate factors until the individual psychologist himself finds significance is surely both arbitrary and subjective. The negative factor loadings (see above) are rotated until they become positive; then "psychological meaning" is attached to them.

The present writer is of the opinion that Alexander's factor loadings, unrotated, mean little more than her own. Rotated, the meaning depends on the investigator.

It would be fairly obvious to assume that our first factor is a general intellective factor. We are not even sure of that, however, unless it be thought of as a kind of functional cluster. It would be reasonable to suppose that our sample of superior subjects might tend to use more "g" than anything else in their performance. We find however that there is a sex difference observable in favour of the boys. It may be, of course, that the girls used mainly 'g' in their performance with a much smaller amount of another special factor or factors, whereas the boys may have relied more on the specific; there would thus be a tendency for the one to cancel out the other, leaving only the combined general factor.

Sex /
Sex Differences.

<table>
<thead>
<tr>
<th>Test</th>
<th>M (boys)</th>
<th>M (girls)</th>
<th>D</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-S. fm bd.</td>
<td>46.66</td>
<td>41.28</td>
<td>5.38</td>
<td>1.75</td>
</tr>
<tr>
<td>Cube Const.</td>
<td>34.6</td>
<td>32.89</td>
<td>1.71</td>
<td></td>
</tr>
<tr>
<td>P.C.C.</td>
<td>29.8</td>
<td>23.64</td>
<td>6.16</td>
<td>1.97</td>
</tr>
<tr>
<td>Healy</td>
<td>69.83</td>
<td>67.32</td>
<td>2.56</td>
<td></td>
</tr>
<tr>
<td>Passalong</td>
<td>55.16</td>
<td>54.66</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Kohs</td>
<td>61.3</td>
<td>54.63</td>
<td>7.12</td>
<td>1.31</td>
</tr>
<tr>
<td>Matrices</td>
<td>49.45</td>
<td>43.62</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Space P.</td>
<td>42.54</td>
<td>33.08</td>
<td>4.46</td>
<td>1.98</td>
</tr>
<tr>
<td>3 Verbal</td>
<td>33.29</td>
<td>33.63</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>Sentences</td>
<td>15.5</td>
<td>17.33</td>
<td>1.33</td>
<td></td>
</tr>
<tr>
<td>Vocab.</td>
<td>17.79</td>
<td>16.29</td>
<td>1.54</td>
<td></td>
</tr>
</tbody>
</table>

The significance of the differences is shown for those tests only which show a tendency towards a significant difference between the boys' and girls' mean scores. As can be seen, four of the tests group themselves together here - Space Perception, P.C.C., Kent-Shakow formboard and Kohs. Since the number of cases in each sample was so small (24 boys) and (24 girls) it is not surprising that \( t \) does not reach significance level. However, there is a decided tendency in that direction and it is reasonable to suppose that with a larger number of cases, the differences in the means would show significance. This is in accord with other findings of a sex difference in favour of boys in performance tests. We have already referred to possible reasons /
reasons for this in another section. The writer is of the opinion that the cause is an environmental one.

The verbal means are almost the same. We did say that we thought sentence completion tests might conceivably be a fairly pure measure of "v", or at least the verbal fluency which, coupled with "g", gives what we think of as "v". It is interesting to note that the mean for sentence completion is the only one of the tests which is in favour of the girls. The difference is not significant, but it would be interesting to see what results would be obtained with a very large sample. We have already mentioned, also, that although we included vocabulary in the verbal battery, vocabulary did not appear to us to be nearly so "verbal" as sentence completion. Again, although by no means significant, we see that the difference is reversed, this time in favour of the boys. With such small numbers these differences may be just due to chance, but again it would be interesting to see what would happen with larger numbers.

What are the tests testing?

This question arises very naturally, especially when we see a group of tests obviously grouping themselves together. Kent-Shakow formboard.

It has been held that formboards are not so similar to other tests involving spatial relationships as we might expect. The reason advanced for this has been the fact that in most formboards the testee is not compelled to "see with his mind's eye" what is required. Koussy (1) reported a subject who wished that the snapes to be rotated were made out of cardboard so /

(1) op. cit. P. 108
so that he could "try them out" without having to make the effort of imaging the shape in the required position. Cardboard could of course be extended to the formboard wooden shapes. Koussy therefore concluded that only performance tests involving spatial visual imagery could be held to be measuring "k".

The Kent-Shakow does not come within the category of boards not requiring some sort of visual spatial ability. The very principle, which we mentioned, of having the insets divided into pieces, prevents the subject from being able to look for complete shapes and he must somehow tend to imagine the necessary combinations at least in order to do the test with a fair degree of speed. This is not to say, however, that the test may not be done in its last stages, that is, where there are many pieces involved for each shape, by a kind of rapid verbalisation of the situation i.e. repeating to oneself what the requirements would be to complete the shapes, supposing that the individual had managed to fit some of the bits in, and was considering verbally what remained to be fitted. Then there would be the individuals incapable of either competent visual space perception, or intelligent inference, who would be at the other end of the scale. With a certain degree of intelligence they could finger the pieces and fit the shapes mainly by the feel of the thing. With low intelligence such an individual would be reduced to pushing the pieces in, hoping that by trial error, some of them would fall together to fill up the required shape. It would follow that the most competent individual with the Kent-Shakow formboard would be one who combined ready visual spatial perception with high intelligence. This is in line with the theory we have already advanced in connection with the individual ways of doing the test P.C.C. It must of course be mentioned that in the first /
first stages of the formboard, in which very few pieces are involved, the most intelligent subjects will fit the pieces with very obviously curved edges first. This can be done quite independently of a factor of visual spatial perception.

It follows that if our theory of visual space perception holds, it must be applicable also to mental defectives. It would be reasonable to suppose that if ability to deal with visual spatial material operates to some extent independently of intelligence, there will be individual differences observable with those defective in intelligence. The writer tested four mental defectives whose ages fell within the same year-group as the large group, that is, they were 13 and not yet 14 years old. The highest I.Q. of the four belonged to a girl who was recorded as 75 (although above 70 she was not considered suitable for ordinary schooling). Her formboard score was 19, giving her an M.A. of 8 according to the norms. The writer was informed that this girl is the best pupil of the Special School. The girl actually told the writer that she liked doing arithmetic best. Now with the formboard, it was fairly plain that she was trying to apply some sort of reasoning to the problem (and with Healy also), only it looked as if her mind reasoned very slowly. So impressed was the writer by this apparent use of reason (not obviously present with the other defectives) that she allowed the girl as much time as she liked to see if she could finish the formboard. She did finish, but beyond the time limit by 6½ minutes. It seemed therefore that this girl could use reason, but unfortunately too slowly to allow her to be educated at the same rate as other girls of her age. Her visual perception with
this material was definitely poor.

Another girl with an I.Q. in the sixties scored only 2, giving an M.A. of 6. Given a very long time she was quite unable to go beyond the fourth stage of Kent-Shakow.

Two defective boys, one with I.Q. recorded as 65 and the other 67, both achieved higher scores with the formboard. The former scored 32 (M.A. 11) and the latter 24 (M.A. 9). The latter was rather unlucky for he fitted the pieces of the first three stages quickly and competently. Plainly he saw quite well, but one piece held him up at stage 4 so that he passed the time limit. The same happened at stages 5 and 6. He was quite incapable of learning by experience (in this case learning that some of the pieces had to be reversed) but he did see the initial spatial 'set-up' very much better than the girls.

It seemed to the writer that the actual sex difference with the material was even more apparent to the observer with the defective children, though no statistical evidence is available with only 4 cases. Obviously the girls cannot, as it were, fall back on "g" to help them out, as children of higher intelligence could do. This, in the writer's view is a very important point. Too often is it assumed that, a certain test being devised to measure a certain factor or factors, must necessarily indicate that the subject is utilizing such factor or factors in the main when he is doing the test. Clearly there are all sorts of different ways of doing tests and it depends on /

* It seems to the writer that an investigation of special abilities with mentally defective subjects would yield very important information, since such individuals have a narrower variety of 'means' to fall back on in solving problems involving special abilities.
on the individual which of his own abilities he utilises in the
test situation. This does not mean that the writer assumes that
a specific test, such as a k-test, does not measure k. She is
of the opinion merely that an individual does not necessarily
utilise much "k" to do the test, although on the other hand there
must be some degree of "k" involved. How are we to differentiate
those with outstanding "k" and average intelligence, from average
"k" coupled with very high intelligence? The writer considers
it likely, too, that girls may tend to make more use of "g" in
a performance test situation than do boys. It may also be
possible that boys use more "g" than "v" in a verbal situation than
do girls.

And before leaving this subject, it would be interesting to
know to what extent the early toy environment of boys and girls
has conditioned what we might call an ability to deal with
spatial relationships. Little boys who are given bricks and
constructive material generally get a much better start in
developing a sense of area than do girls, confined as they are
in the main to dolls. Given the requisite level of intelligence,
we might be able to increase the number of female mathematicians
by some careful toy-conditioning! We are of course neglecting,
for the moment, the question of differences in a special ability
to deal with spatial relationships being innate. It looks as
if we must put this aside completely until possible environmental
causes are thoroughly explored. The writer inclines to the
"environment" view, but of course the possibility of the other
must be kept in mind. Also it is feasible to assume that the
end result may come from predominating environmental influences
plus /
plus certain innate factors, the latter determining in part one's type of visual space perception for instance. Obviously different people rely on different cues. It is possible, for instance, that kinaesthetic cues are also a type, or a way of solving a certain orientation problem. It is possible, too, that a problem in orientation is most quickly and therefore "best" solved on the basis of kinaesthetic cues. That is not to say, however, that it may not be solved another way, according to the cognitive means at the disposal of the individual solver.

Kohs blocks.

There appeared to be at least two different ways of doing this test. While the writer watched, she characterised them mentally as the aesthetic method and the analytic method. Perhaps "direct" would have been better than aesthetic for the aesthetic method was part of it only, and occurred especially with children good at art, who saw the aesthetic and maybe the form relationships very readily. At any rate, when done by this immediate method, in which each design is built up, adding as it were to the coloured 'pattern - Gestalt' without really thinking of the blocks as separate, the whole test tends to be done fairly quickly. On the other hand, individuals who used this method sometimes got muddled up, and where their form perception was not good, tended to over estimate the relative sizes of the diagram and the blocks available. This tended to give a characteristic result with the diagram involving a yellow diamond on a blue background. The subject had no idea of estimating the number of cubes necessary for completing the design, and often made the inside yellow diamond far too big.

When /
When the last two diagrams of the $10^{(1)}$ are given, such a subject gets very muddled up, for there is no immediately apparent design to build up. When this "whole" approach is, as it were, combined with "good seeing", the subject achieves a very high score. This would be the combination of analysis and synthesis which Kohs himself mentions as characteristic of the test $^{(2)}$. Again this is probably the quickest way of doing the test and the "best", but it is not the only way and it does not necessarily depend entirely on intelligence, for with our selected group there was a very obvious variety of methods.

What we called the analytic method was very much a question of dividing off the diagram mentally into cubes. The very competent individuals using this method seemed actually to build bit by bit, very much aware of each separate bit as playing a part in the whole. Such an individual might start off with the top corners and work down the pattern in blocks of four, as would be the case when 16 cubes were given. Using this method, the individual's score will depend on the extent to which he uses the method competently. On the whole this method is more time-consuming than the "total" approach. And as we saw, the test is scored in our case on the basis of time. Counting the number of moves would not differentiate between the methods, on the other hand, for the number of moves might in two individuals be approximately the same by using two quite different methods. The only way to differentiate would be rate of moves, that is actual /

---

(1) This is the Dreyer & Collins selection from the original 18 of Kohs.

(2) "Intelligence Measurement". 1933.
actual time taken altogether.

This impression of the writer of two quite distinct methods was confirmed by the subjects' introspections. She briefly explained to each subject the possibility of doing the test in certain ways and asked each one to try to describe his particular one. Some reported a mixing of the methods, which we of course might quite expect.

The possibilities of the Kohs blocks test have not been nearly widely enough exploited, in the present investigator's view. Even from the point of view of differences in individual approach to the test, it would yield very interesting information. The actual test material itself, also, could be much extended.

Healy Picture Completion II.

Again two ways of doing this test were noticed by the investigator, and there may be others. She asked each subject for the reasons for each choice in completing the pictures. She found that subjects tended to rely either on direct observation or on imaginative reasoning. For instance there were two ways of finding out what the boy in the picture was throwing—either you noticed that he had a blue cap on the nail in a previous picture and therefore assumed that being bare-headed this blue cap was missing; or you wondered what the boy could possibly be throwing so mischievously and thinking back maybe to personal or observed experience might conclude that it was some personal /

* In connection with the question of methods relative to the individual - one girl who gave an outstanding performance on Kohs was found to be extremely good in art. Compare her high Kohs score with the same score achieved by an individual using the analytic method to a very competent degree. To say that those two scores mean the same psychologically is obviously untrue. Yet this situation must occur in many tests.
personal article such as a cap. Or again, either you concluded that the boy had dropped a red book from his pile in another picture because you had noticed that in a previous picture his pile of school-books lay on the table, the top book being red; or else you estimated that the distance between the loose strap and the remaining books was enough to hold one book and therefore you looked for one book. This was probably why some people chose a blue book instead of a red one, since not having observed the exact colour in a previous picture, any single book would do to fill up the gap. And we can only guess at the extent of lack of observation in the individual who does not even notice the precariously swinging books within the strap of the walking boy, but imagines that a white cat following after makes a good picture. A mentally defective child may also choose the white cat, as happened with one of the members of our small defective group: Was this imaginativeness, or just lack of reason?

There were thus two obviously clear methods of doing Healy-direct observation or imaginative story-completion. A high score might be achieved by either method or by a combination. Again, identical scores may mean psychologically different things.

We can therefore see from this short description of some of the tests that we must find out just exactly what these tests are testing and relative to each individual. The time has come, it would seem, to stop putting together batteries of performance tests and assume that they are testing intelligence, or an "f", or practical factor. A thorough investigation of single tests would be very useful. Kohs, for instance, would lend
lend itself admirably to this type of analysis.

We have seen, too, to what extent special abilities may be operating in test results, and how uncertain within those results the relative predominance of special ability over intelligence or vice versa may be.

Finally we may assume that an extensive special disability would certainly tend to disturb the integration of temperament, at any rate with school children. We should expect to see evidence of such disturbance in the actual quality of performance in a wide test situation.

It is thus because of this combination of possibilities of investigating intelligence, special abilities and temperament in a "total" approach, that we consider the performance test situation so valuable.
CHAPTER VII. METHOD OF TEMPERAMENT ASSESSMENT.

The behaviour of the subjects during the performance test situation was to be the sole standard by which we decided to try to make some kind of assessment of temperament.

The writer was aware that there would be two aspects of each performance — an intellectual and a motor aspect. The former would be revealed in degree of planfulness exhibited, presence or absence of trial and error, care for detail, the reaction of the subject to a situation of difficulty, and the fluctuations in type of work. What would be revealed on the motor side was not so obvious to the investigator. She therefore drew up a list of points under the headings of Movement; Speed; Inhibition; Type of Activity (e.g. deliberate, systematic, inaccurate, confused etc.); Shape; Method (reasons, logical or aesthetic judgment, tries pieces at random etc); Persistence of effort; Use of Hands. This list was therefore a mixture of the two aspects we have mentioned. On each point every subject was rated on a scale ranging from +3 to -3. This rating was of course not done during the actual test but as soon afterwards as possible. During the test the investigator made a few notes in as unobtrusive a way as possible. It is, of course, obvious that all those ratings were/

* All the performance testing was carried out in the afternoons. Ordinarily, the morning is considered to be more favourable from the point of view of freshness, but the afternoon suited both tester and school best, so that even where a morning session was possible it was not taken advantage of, so that conditions for all subjects should be similar. The group tests were all given in the mornings, on the other hand.
were subjective, and therefore relative to the writer's own estimate of each pupil. She hoped, however, that out of this mass of rating something might possibly emerge which could be used in a relatively objective way at least, that is to say some point or points which other investigators might utilise which could be to some extent independent of the individual rater. Only in this way could some sort of basis for an objective assessment of temperament be laid. And even if some such point did seem to emerge it could only be validated on the basis of agreement with an independent rating of a reliable and constant observer, such as a school-teacher.

It was decided that the profile method was the best way of representing the individual traits, so on the basis of her own estimates, the writer was able to graph a profile for each subject tested. These profiles were drawn on the basis of what appeared to the investigator, fairly well marked, traits.

In order best to describe the results we shall divide the subject matter into two parts, the first being a description of the traits observed, and the second being an account of an attempt to validate an aspect of the temperament assessment by comparison with teachers' estimates and including the use of the chi-square test of significance of relationship.

general ratings.

From the data gathered about each individual it was decided that the following general traits, when taken together, could give a description of a subject which might usefully represent his temperament. These traits are speed, smoothness, flexibility, Speed/
Speed of Decision, Use of Hands, Freedom from Load, Confidence, Motor Inhibition, Care for Detail and Sex Dominance. It must be emphasised that these traits are meaningful only with reference to the complete profile, and by themselves do not signify very much. Some explanation of each must be given.

Speed.

This was an estimate of the actual speed of movement throughout the tests. It is obvious enough what we mean by this and requires little elaboration. It ranged from the extreme of excited, restless movement in which there seemed to be a general exaggeration of each movement necessary. The pieces in the formboard were rapidly put in place successfully if the subject was competent in the visuo-spatial relationships, but if deficient in the latter this type of over-movement led to many useless and inefficient moves, often including the pushing and snapping of the pieces against the sides of the formboard shapes. This kind of urgency also led the blocks in Kohs to be pushed too far, perhaps upsetting what already had been completed of the pattern. (Usually, however, the extreme of speed was not present in a subject very competent with the performance test material). There seemed also to be wastage of effort, and when the position of a block was changed, such as in Kohs or in Cube Construction, it was often moved through too great an angle, so that inevitably half of the movement was entirely useless. An excess of speed with a test like P.C.C. tended to lead to trial and error tactics in the constructional part of the test, which in itself necessitates a rather measured, careful rate. Often an individual of the extremely speedy type scores higher than a slow subject, or even higher/
higher than the average type of reaction, because sheer number of movements must produce a certain percentage of efficient reactions. The score will be higher in proportion to the level of I.Q.

Speed thus ranges from this extreme type through the normal, where there may be mild signs of excitability at the beginning, soon quietening down, however, so that the theoretical normal should show neither too much nor too little speed, and in accordance with degree of intelligence, adapting well to each novel situation.

The other extreme would be manifested by extremely slow moves, even errors being repeated with the characteristic lack of speediness. Although not present in any number of an selected group, we can imagine that the real extreme of this kind of reaction would be the tendency to apathy which may be in evidence with the very feeble-minded and the psychopathic.

The assessment, as in all the ratings, was from +3 to -3, the normal type of speed being assessed as zero. The distribution within the group was 13% tending to the slow end, 62% in the middle and 25% tending to the speedy end.

Smoothness.

This may seem to be a very vague and undefined trait. It is really a type of movement, and before the actual performance testing, had never occurred to the writer as a possible "item" of a temperament schedule. However, it was so "real" during the behaviour reactions of the subjects, that the investigator felt obliged to record its presence and rate each subject in its light. It was really striking the extent to which the population could be divided up into smooth moves, jerky moves and the large middle part of/
of the population who were neither the one nor the other to any marked degree. At first the writer thought that above-average smoothness might be correlated with above-average artistic sense. So sure was she of this that she asked each "smooth" reactor as he came along whether or not he was musical, played a musical instrument, or liked to draw or paint. It was strikingly characteristic, also, of one individual with the curved palm and pointed fingers, so often associated with an artistic hand. This particular boy we are referring to, reported that he played the piano and practised regularly, that in addition, he liked to paint in his spare time and that his father painted a great deal. The smoothness of his handling of all the material was so striking, that his performance in general seemed to be a series of very pleasing curves. This connection with aesthetism, however, did not always hold, as the investigator gradually began to realise! She found it also occurring in individuals who reported no aesthetic interests. Also, she suddenly realised that a friend of her own, who was a very good pianist, exhibited in many ways this very jerkiness which she was beginning to think was characteristic of the non-aesthetic.

There seemed to be nothing else to correlate 'smoothness' with, nothing at any rate that occurred to the writer. But, as we have already remarked, its presence was so real, and it appeared to be something so fundamental and generalised in the individual, that it had to be noted down.

In re-reading Kretschmer (1) we found that this was very smooth-jerky bipolarity occurred as one of the typical ways of differentiating the cyclothyme from the schizothyme. The peculiar jerkiness/

(1) Op. cit. P.
jerkiness of all psycho-neural processes was held to be characteristic of the schizothyme, as compared with the smoothness of the cyclothyme.

We also find C.J.C. Earl remarking on the same trait. He says "The nature of the rhythm of psycho-neural process may be the most importance of all differences between personalities, for all temperamental divergencies may prove to be dependent upon the smoothness or jerkiness with which the pattern of excitation spreads in the nervous system.

flexibility.

This trait-name is really self-explanatory. We mean the ability or lack of ability as the case may be, to change readily from one thing to another. What we intend to signify is not really perseveration. It may be influenced by an innate tendency to perseverate or not to, but we mean to imply some degree of volition, so that what we are estimating is not intended to be a pure measure of function, secondary or primary. It may be objected, probably with considerable justification, that a performance test situation is not nearly varied enough to give enough insight into degree of flexibility. This may possibly be true, but the investigator did see that some individuals, as each new test was presented very readily set to, immediately after the instructions, as if the previous test or tests had not existed. At the other end of the scale were the subjects who did not immediately "bite into" the new test, but looked at the material as if it were a wrenching into a new situation. They were not necessarily people who were cautious, or who paid over-much attention/

attention to detail. With Healy, for instance, there might be preoccupation with the actual pieces rather than with the picture on the piece, and the tester might be asked if they were meant to fit the squares exactly. Again the subjects appeared to be distributed fairly normally on this trait. 64% of the group fell within the normal range, with 16% on the non-flexible side, and 20% on the flexible side.

The writer believes that a more varied test situation could quite easily be devised in which this trait could be more clearly estimated, although in her own estimation it was so demonstrated with the test battery even as it stands.

**Speed of Decision.**

Again this trait-name is self explanatory. It signifies the speed with which a subject makes up his mind. At its extreme end, the individual makes little use of diagrams when available, as in P.C.O. or in Kohs, or Passalong. Especially was it to be observed with Healy. At the other extreme the individual makes laborious use of diagrams, consulting them repeatedly and with elaborate care; he also tends to show "second-thoughts" with the Healy choices and waves with one piece in poised hand and with eyes sweeping the other possible choices in the box, sometimes over and over again. Once more the majority of the group fell in the middle portion of the range. 78% fell within the middle portion, 22% within the slowly deciding group and 8% within the speedy group.

**Use of Hands.***

Of all the psycho-motor behaviour which attracted the attention of/

* represented on the "profile" by 2 circles for 2 hands, one red circle for right predominance, and one blue for left predominance.
of the writer, the use made of hands in manipulating the material was by far the most striking. There was a clear and distinct range from the entirely unnecessary, elaborate use of both hands through the middle range of predominant use of both, but in a merely adequate way and no more, or else predominance of one hand while helping with the other, to the other extreme of the one hand being used throughout the test while the unused arm was kept stiffly at subject's side, often with the hand clenched and thumb turned under the fingers.

The first thought that occurs to mind is the handed-ness of the subject — whether or not so and so was left or right-handed, and whether or not somebody else might be ambidextrous. The writer is of the opinion that in the special situation described these facts have but a superficial bearing on something much more fundamental. She tested the eye-dominance of each subject by the simple test of rolling a cylinder of paper and asking subject to look through at some distant corner of the room, and whenever there was a predominant use of a left hand she investigated such a use afterwards. An odd bit of information, sometimes emerged, such as the boy who used mainly the left hand while helping slightly with the right, for the performance tests, and informed the tester that he wrote with the right hand, but threw and kicked a ball with left hand and foot respectively. He reported that he had been changed over to the right hand by a primary school teacher. Such pieces of information were interesting, but did not really change what was the important part, namely that what mattered was whether the subject tended to use only one hand, or a mixture, or predominantly two. It did not really matter whether it was right or left hand — what seemed important was the singleness or double-ness of the approach. +3 and -3/
-3 were allotted for each extreme, and a zero score was given either for the use of both hands with poise and discretion or the predominant use of one, helped adequately by the other, both hands being well in sight of the tester. The group were distributed as follows - 65% filled the middle part, with 23% tending to both hands and 12% to one hand. At first glance it would appear that we have found just one more manifestation of the old dichotomy, extraversion-introversion, translated into terms of movement. However, as we shall explain later, we mean to imply rather more than that.

There is one further possible objection to this trait which we can fore-see; that is the fact that the use of hands may be one of those personal idiosyncrasies to which no great psychological importance can be attached, or at any rate no more than individual importance. Our reply to this would be that in the performance test situation, and with the formboard in particular, the subject is compelled to use his hands. It is not just a case of gesticulating or making side-movements as the subject feels inclined, nor is it a case of using hands when some other action might do; instead the only way in which thought, and the main way in which temperament is expressed, in the overt action necessary with the performance test material must be through the hands. There may be other motor manifestations, such as facial expressions, but those are not really necessary for coping with the situation except relative to the individual. The hands are so necessary, and the writer decided that not only in terms of the individual profile, but in a generalized way was she justified in picking out this trail.

Freedom/
Freedom from Load.

This also was something which had not occurred to the tester as a possible trait. She has taken over the name from Downey (1) but the actual trait itself she observed quite independently. Downey's name for the trait was however a better description than the one the writer thought of using. Downey measured it by the ratio of natural speed to capacity speed. The present writer, however, merely noticed that some individuals appeared to need a warming up period before each test and that at the extreme end of the scale such individuals never really get started at all. This may either be connected with perseveration or with an excessive tendency to relax. Downey describes this trait very adequately "Freedom from inertia suggests that the psychic machinery runs with little friction; it is not necessary to oil it continually in order to keep it going smoothly and speedily; its own inherent energy holds it at top speed".

The estimation of this trait was necessarily the most subjective of all the estimates, but however much it depended upon what the writer estimated as average, above average and below average, and however ill-defined the trait may seem to be, the investigator felt impelled to record it. The distribution ranged from 72% in the middle, with 8% showing load, and 20% freedom from load.

Confidence.

Again we have a more or less self-explanatory term. The greatest/

(1) Op. cit. Downey's final term was "freedom from inertia". "Load" was the first way she described it. The term "load", however, suits the present writer better.
greatest danger in estimating this trait will be its interpretation in terms of the individual tester. It can be argued that the estimate of a child's confidence or assurance will depend on the extent of some trait in the psychologist. This is undoubtedly true, but we would add that nobody who exhibits either an excess or extreme defect in this respect should be a mental tester. Surely some sort of common ground can be established in which a trait like this, which should be useful in a vocational guidance interview, could be assessed. For instance it should be possible to make cine-camera records of a large sample of the child population at different ages and establish thus some kind of norm of behaviour in terms of this trait, and of others too. For psychological testing, it seems to the writer, such cine-camera records would be of inestimable value and would probably be a more fruitful source than any other for developing norms of behaviour. If any characteristic patterns were observable, they ought to become so, should such records be observed often enough and long enough.

The main difficulty with regard to the estimate of general confidence is not so much the individual psychologist, although that is doubtless a big enough difficulty by itself, but the wider differences of social groups and nationality. Obviously it would be necessary to establish separate sets of names to begin with, although this would have to be done with tact and discretion; otherwise it might lead to individual results. In one group the distribution ranged from 18% lacking in assurance, 63% filling the middle portion, and 21% at the "assured" end of the distribution. Possibly more than any other trait, also, this one would be subject to fluctuations from one test situation to another. It is very much influenced in the test situation also by special abilities and intelligence.
Motor Inhibition.

The other extreme of what Downey called motor impulsion, is what we mean by motor inhibition. Again this trait was chosen quite independently of Downey. It appeared to characterize a certain type of reaction within our own group which, it seemed, by including in our list of traits, would help to interpret the individual profile more insightfully.

The freedom of the subject's energy was determined by the degree of motor inhibition. When there was the smallest degree of such inhibition we formed the explosive type, who not only was very free with movement and tended to be speedy, but broke out into spontaneous conversation. (We do not include here of course the occasional talkative individual who by parts of his conversation reveals that the actual talkativeness is conditioned by anxiety).

At the other extreme of the distribution we find the highly obstructed type of individual. But here we must make a subdivision into the subject who shows marked motor inhibition, giving the impression of pent-up nervous impulse which, given some special stimulus, would burst forth with just as much energy as the explosive uninhibited type, and the subject on the other hand who exhibits the same poverty of reaction as the former individual, but is actually "nerveless" and indifferent. What we are describing corresponds really to the two sub-divisions of James' type of obstructed will, which he defines in terms of two possible types - exaggerated inhibition (our first example) and deficient impulsion (on second type).

The/

The extreme type of motor inhibition is exhibited in a general under-movement with the performance tests. Adequate use never seems to be made of the individual's intelligence. The impression is given that the intellect is pent up in some way and that the subject is unable to adapt sufficiently freely to meet the needs of the test situation. Especially is this inhibition in evidence at a moment of difficulty, for instance at the last two stages of the Kent-Shakow formboard and at the beginning of the constructional part of P.C.C. In the latter test even where a subject has the special ability which is plainly necessary to complete this test within the time limit, a subject of high inhibition may have a bit of the wood actually turned round the correct way but may make so tentative a move towards completing the step, that he fails to score in the time. The writer has one subject actually in mind, who very clearly saw what was required at step VII of P.C.C. for the wood was held in the correct position, but inhibition was so great that only the most limited attempt was made to carry the thought into action. The necessary success in overt action was therefore not attained. An individual of a very much lesser degree of intellectual ability, but temperamentally tending towards the other extreme of this trait might through sheer force of trial and error and accumulation of movement stumble on the correct solution. If the degree of intelligence in this latter subject were high enough, sufficient learning might take place to bring about successful overt action.

Earl, in discussing the performance test results of morons, describes/
describes this same inhibition reaction-type and considers it to be more serious than the excitable type. When it is very marked, or when the other deviation of apathy is present, he finds it characteristic only in definite psycho pattern of schizoid or hysterical colouring. With our own group no such extreme case was found but there were certainly two subjects with a slight tendency in that direction.

Care for Detail.

The extreme pole of the trait would be exhibited in excessive pains to be accurate, giving the tester an impression of what might amount to an obsession. It seemed to the writer also, that this was one of the most important traits on the profiles for an individual manifesting such extreme interest in detail seemed thus to colour his whole personality with one trait. It is important that it should not be confused with logical keenness, which is more a matter of the intelligence level. We are not, of course, supposing that care for detail can be separated in any absolute sense from the individual level of intelligence, but we certainly do stress the need for keeping the two attitudes separate. If care for detail excessively predominates the individual personality, we have the kind of subject so absorbed in the unimportance, that he cannot see the wood for the trees. This is well illustrated with the Kent-Shakow formboard with the latter stages especially where many pieces are involved. This type of attention to detail will compel a less intelligent subject carefully to try pieces in all sorts of different ways, using perhaps the same few pieces over and over again, and totally neglecting much more obvious pieces which are yet to be inserted, such/
such as the very curved slopes. This is the type of subject, also, who gets so swamped by detail that he tends to be slightly overcome when there is a large variety of choice as in the Healy picture completion. The higher the intelligence level, the better will the subject be able to differentiate the general from the particular, but even with high level of I.Q. there may be some slowing down in output.

The distribution in our group from the extreme interest in detail to the other pole of carelessness and leaving unfinished for a time, because of diverted attention, some quite easy to complete shape, e.g. in the formboard, or the acceptance of hastily placed pieces in the first stages of P.C.C., was as follows - 48% in the middle, "normal" position of the trait distribution, 22% tending to the care for detail end, and 10% tending to the careless pole.

Sex Dominance.

The reason for including sex dominance in our list was not that we were able to assess it so exactly on test performance, but because we realised that certain types of temperament profiles might give a slightly misleading account of the individual with reference to the general level of the profile. Some of the traits which we have mentioned are negatively correlated, so that it is uncommon to find a uniformly high profile. This, however, may occasionally happen, when the subject has what we might call specially positive temperamental traits. A fair degree of motor inhibition may quite possibly be combined with quite high flexibility for instance. It all depends on the individual's total personality how those traits are interpreted. A certain control of inhibition, the power to keep
an impulse under control, is essential in most occupations. A certain amount of load, coupled with attention to detail will be useful for the painstaking scientist, but as Downey points out, it is probably flexibility which is the saving grace of the more creative scientific worker. We are, however, contemplating no scheme of connecting up our traits with either types of adult worker or with types of occupation. We have no adequate or reliable enough basis for such elaboration, and in any case such speculation would introduce a premature rigidity into our concepts. The writer would feel compelled to carry out a very long and expensive piece of research before she could believe herself justified in any such prediction.

But we must return to sex dominance for a moment. If a subject did come out high on several traits, such a profile tends to give the impression of a rather strong and almost aggressive personality. This may be quite misleading, for one boy who did achieve such a profile had a distinctly feminine element in his personality, so that we were led to estimate normal for either sex as 0, and in this case, we assessed the subject as -1, that is slightly below average in "maleness". It seemed to the writer that this brought his profile into proportion. A -1 score for a girl, similarly would mean a slight lack of femininity.

Obviously with boys and girls of 13 there are few particularly outstanding secondary sex characteristics which we might assess exactly. What we did go by was delicacy or lack of it in handling the material, texture of skin, and general physique.

No idea of such an estimate had occurred to us at the beginning of the testing, but as we went on, we realised more and more that something/
something had to be brought in, to balance an occasionally misleading profile. The dominance of masculinity or femininity appeared to the writer to be adequate.

In a later section we describe some individual profiles in detail and see how the total estimate compares with their general scores, school achievement and teachers' estimates.

The writer decided that to work out the correlations between the traits served no useful purpose. Any consistency which might emerge would after all be a function of her own rating and would have no sort of meaning apart from such internal relationship. Also the estimation of correlation coefficients is not really applicable with the particular system of scoring used. To have used the chi-square test with the ten traits would have been both laborious and of doubtful value. There was after all no outside criterion with which to compare the tester's ratings. As we shall explain in the next chapter, the comparison of ratings was narrowed down to a smaller field for several reasons.

Special Reactions.

The investigator was struck by certain reactions whose significance was not quite clear. The first of these was (a) several pieces held in the hand, that is pieces of the formboard which the subject had picked up, had not been able to place correctly, yet held them in the unused hand. The same applied to any pieces of P.C.C. which were not put down again after an attempted move, and it was possible to see this also with Healy picture completion and Kohs. with the last two it was, however, less in evidence than with the first two.

The first time this reaction attracted the writer's attention was during/
during the try-out of P.C.C. with the unselected group. She had observed it once or twice before she realised that it was likely to mean anything. Before the group was completed, she noticed this clinging to rejected pieces in two further pupils. On investigation, they were both found to be rather competitive in school. When this reaction occurred again in a very few cases with the selected group, it was not so clear that the pupils were competitive. In fact one boy does very poorly in class relative to the other pupils. He did appear to the writer, however, to be very inhibited, and wore at times rather an ominous and aggressive look on his face.

It is interesting to note that Earl(1) also observed this reaction and did not quite know with what it was related.

(b) Thumb held under fingers. The writer could not quite make up her mind whether this was a reaction of anxiety, or whether it was some kind of attempt on the subject's part to brace himself for action. It certainly was quite a frequent response (or reflex), and was typical, at any rate in one case, of a generally anxious boy. The latter subject was over-talkative, described how he spent all his spare time making models of boats for his much younger brother, how he suffered from rheumatism after a bout of rheumatic fever, and how he quite often found himself standing up in bed in the middle of the night as a result of nightmares. It looked as if the boy were being made to assume a great deal of responsibility for this younger brother, and in his case the writer/

writer could not help wondering if the thumb position were not some kind of infantile regression during a position of difficulty. This subject also tended to watch the tester very closely.

Dr R. Rusk (1) informed the writer that this reaction had, a long time ago, been reported to by Warner in "The Study of Children". It was described as a position of the hand in fatigue. This may be so in some cases, but the present investigator observed it right at the beginning of her test battery, sometimes, when fatigue was unlikely to be present, and at other times at the beginning of a new sub-test in which the individual seemed, by it, to indicate a renewal of effort.

It seems at any rate to be connected with some kind of nervous tension, and did not in general come afterwards (as surely would be the case, with fatigue) but often before some new piece of activity. F.L. Golla and S. Antonovitch (2) found that the increase in tension associated with the performance of an intellectual task actually precedes the commencement of the intellectual activity. This fact led them to propound the interesting theory that the increase is an atavistic remnant, a phylogenetic relic of a reaction preparatory to physical action.

(c) Rigid clumsiness.

Only twice throughout the performance testing did this reaction appear. In both cases, the subjects sat with body very rigid, and with mouths very tightly pulled in. This air of clumsiness was most apparent with the formboard. The pieces were dashed in, with

(1) Glasgow University Educ. Dept.

a clattering, exaggerated noise. One subject in particular gave other signs of anxiety and insecurity in frequent and very audible signs of which he himself seemed totally unaware. An occasional tremor was also visible at the corners of his mouth.

The evidential value of this reaction was not at all clear. It may have been a simple anxiety reaction or it may have been due to an inherent organic factor or factors.
VIII. ESTIMATION OF A GENERAL TENDENCY OF TEMPERAMENT.

In our discussion of single traits we indicated that what struck us most forcibly in the entire performance test situation was the use made of hands. Although the writer isolated what appeared to her to be a fairly useful selection of traits, the degree to which each was present in the subjects being represented by means of an individual profile, the tendency which appeared to her to be most real and most constant and most likely to be individually typical, apart from the single test situation, was this single or dual or mixed handedness.

We have already indicated why we consider this trait to be a likely generalised reaction. We must now amplify that point of view.

It has all along been our contention that the chief virtue of any particular theory of temperament which we mentioned favourably was the possibility of its being dynamic, and its applicability in a "total approach" to temperament. When we talk in terms of totality, we do not of course imply that the whole is indivisible into parts, but merely that the significance of any trait or traits measured must be in terms of their relevance to the total personality. We never at any time intended to survey the whole field of temperament; a much more extensive and far-reaching period of research would be necessary for that. But we did have a slight hope that somehow or other we might be able to tap the temperament totality to the extent of measuring some part of it which, although an integral part of the "totality", might yet permit of a certain degree of isolation.

Our/
Our purpose in selecting the traits which are represented in profile form was therefore but a half-way stage in our investigation. From amongst those traits we felt compelled to choose as the principle variable the individual use of hands. We have already discussed the range of assessment. Hands alone were not enough, however. For there were individual differences in rate of movement as well which seemed, if combined with the former trait, to throw more light on what we decided to call Activity. We purposely do not designate it "motor"activity, for we believe that it can be regarded in a generalised way. After all, an individual may manifest a very high degree of intellectual activity, but may by no means be what one calls an active man in any complete sense. On the other hand, the person who is always "on the move" may display but little mental activity in a sustained sense, although sometimes his mind may be active in a "dispersed" manner. What we are trying to say is that our general trait of activity can only be interpreted in each individual relative to his intelligence level, relative to his special abilities, and relative to his emotional range. As well, it indicates potential, in addition to actual activity, we believe. Possibly also, if we could only measure it reliably, we should include primary or secondary function. A general trait seems meaningless to the present writer, unless it can be so interpreted. The same approximate level of activity in two people may have entirely different significance for each. In one it may be merely fussy, apparently pointless busy-ness, which may serve no really useful end, while in the other, if combined with superior intelligence, such a combination may lead to most fruitful endeavour. Activity by itself may be misleading.
malingering, but activity plus method used in the tests, can give us a meaningful picture. Only, therefore, in terms of other measurements can we interpret this variable, but we believe that thus interpreted it can prove to be of some value.

We mentioned very much earlier in this investigation, our preliminary use of a short battery of performance tests with adult subjects. We found that their intellectual behaviour was very much of a pattern - there was great planfulness, precision, and a high level of achievement. All the subjects were personally known to the writer, but at the time, although she seemed to see differences in their actual motor behaviour, and although, their introspections told of emotional differences she could not quite pin down anything definite. At the end of her investigation with the selected group, she decided that if this combination of hands + rate was a real temperament indicator, she ought to be able to predict, on the basis of her personal knowledge of the adult subjects, how they would react with another performance test. These subjects had not done the Kent-Shakow formboard, so it was decided to try them out with it. Only two of the original subjects were available, but fortunately they showed distinctive differences in temperament to a much greater extent than intellectually. The one is, generally speaking, inactive in nearly every way except in a highly specialized intellectual field in which he happens to be on a first-rate level. What he produces intellectually, however, despite the many characteristics which the man-in-the-street would term "lazy"/
"lazy" far outweighs in value the vastly higher level of activity of many other individuals, who remain but conscientious plodders. If his general activity were more intense, on the other hand, a greater volume of valuable work would result. Accompanying his inactivity there is inevitably, some inhibition. The writer therefore predicted in advance that this subject would use one hand throughout the whole formboard. This was precisely what happened. We cannot urge strongly enough, however, that this must be interpreted relative to the individual's intellectual performance, for the latter was in every respect competent. There was no trial and error, the pieces were placed in a slowish measured fashion, and there was every indication of ability to deal with spatial relationships. Each stage of the board was completed well ahead of the time limit.

The second subject, although his intellectual performance was on a similar high level in the other tests given a year earlier, yet appeared at the time to the writer to exhibit some sort of difference in motor behaviour which she could not define. Now, in the light of her knowledge of the subject's temperament and the results of the group testing she again could not quite make up her mind about him, for combined with a fair degree of inhibition he yet exhibits a certain manic aspect in his make-up. He is a much busier and more active person altogether than/

* A certain degree of inhibition is probably a "good" and inevitable sign in workers in intellectual fields who must so often suspend judgment and not jump to hasty conclusions. It probably indicates the necessary control of impulse which must to some extent be a necessity.
than the other subject. Because of the inhibition which at
one time was probably stronger than it is today, the writer
felt that there would be a tendency to use one hand only. On
the other hand, his sheer activity would indicate dual manipula-
tion. When he was given the Kent-Shakow formboard, he started
out with the right hand only, but in a very short time began to
help with the other one, and by the time he was half-way through
the test both hands were in use in a capable and poised fashion.
His actual rate of movement was also greater than the first
subject's although it was by no means too great, but merely the
expression of a quicker and more active person. Again the
writer's prediction was fulfilled. The interesting point, too,
was that both subjects were sophisticated as far as test material
and the test situation were concerned, but each gave a highly
individual motor reaction.

Finally the writer decided to give the Kent-Shakow formboard
to a young boy with a fair degree of visual spatial perception,
an average level of intelligence, and a very "active" temperament.
He is the type of person who takes part in many different spare
time pursuits and has always got some project or other "on hand".
Were he a woman, he might possibly be a meticulous and rather
fussy housewife. It was predicted that he would make lavish
use of both hands, yet at the same time doing the formboard
competently enough because of his special ability. Again the
preliminary prediction was fulfilled. Both hands were used
right from the beginning, the eyes of the subject moved freely
over/
over the board, and the actual speed was above average.

The investigator therefore felt reassured in her theory that she had found at last what was differentiating the motor performances, when the intellectual behaviour had been about level. Apart from the vast influence exerted by environment, the writer is inclined to be of the opinion that this generalized activity trait is fundamental to each individual and may therefore be regarded as a "real" temperament trait.

We may recall for a moment the three variables which Heymans made the basis of his temperament theory. They were Emotionality, Function (Primary or Secondary), and Activity. The various combinations of those, he later differentiated into types.

We remarked, in describing this theory, that it seemed to be a good basis for temperament assessment because of its dynamic quality. No estimate of temperament can be made without some kind of measure of emotionality. The writer thought she could distinguish certain emotional differences in her selected group even with performance test material alone, but she was also well aware that this was by no means an adequate enough situation for any kind of measure of emotionality other than an extremely rough one.

Activity, on the other hand, she believes was open to measurement. Not only that, but such a basis of measurement fulfilled the initial stipulation of objectivity. Surely any tester, after all, can tell whether a subject has used one hand all through a performance test or tests, and can see (as happened with/)

(1) Op. cit., P. 15
with one or two extreme cases in our group) whether or not the other arm is held rigidly by the subject's side. Through the wide middle range we find predominant use of one hand or the other, or mixed hands. Tending towards the other pole we find more and more elaborate use of both hands, carrying such manipulation to an unnecessarily lavish extreme. And the type of manipulation can fairly easily be interpreted relative to the speed of movement. For instance there might be adequate use of both hands, but the actual speed might be slowish. This would therefore reduce the score somewhat on a positive side of the curve. Similarly the use of one hand predominantly, but coupled with fair speed would increase a negative score to some extent. The point scale chosen might vary. (In our own case we rated from -3 to +3). Also, if funds became available cine-camera records might be made to establish norms of reaction.

Comparison of Trait with Teachers' Estimates.

We have already explained that we did not consider it useful to ask teachers to rate the pupils on the isolated traits of our profile. We had several reasons for this decision. They were as follows -

(a) rating all 46 pupils on 10 traits each would be too lengthy a piece of work to ask any busy teacher to do. The sheer number of estimations necessary would inevitably lead to a certain carelessness in assessment.

(b) it was by no means likely that the traits would mean the same thing to the two teachers who were going to do the rating and in any case they might tend to rate in terms of the pupils' achievement in their particular school subject.
(c) the teachers were concerned with secondary-school subjects and were not to be expected to be so familiar with the pupils that they could accurately gauge each pupil's emotional reactions or even the individual variations in intellectual method.

It was therefore decided, that not only was the activity variable the only one which might be real and generalized in the sense of innate temperament, but also, school teachers were much more likely to assess general activity more reliably than emotional reactions, with which they might not feel themselves concerned.

We made up our mind also, that the best way to have our variable rated was in verbal terms rather than in terms of pure symbols. We thus drew up the following scale, giving an actual range of +4 through zero to -4 but verbally obvious at the same time.

| Inert and Listless | Calm and Deliberate | Normal degree of Activity | Quick and Vivacious | Excited, restless, Unable to keep still |

It was explained that one point should be turned into a cross to describe the individual pupil and that if the individual fell between two descriptions, to turn the line into a cross halfway between the two points. One class teacher rated 31 of the cases, and a class teacher of the remaining 17 did the second rating. One teacher could not do it all, for the subjects tested were/
were taught in two separate sections, although both were in the A division of their year.

Calculation of Significance.

Obviously the \( \chi^2 \) technique was the best statistical method to employ in finding the degree of association between the two ratings.

The scores of the teachers ranged from +4 to -4, and those of the writer from +3 to -3. It was decided to divide the case into \( 2 \times 2 = 4 \) categories according to whether they had positive or negative ratings from investigator or teacher. This procedure is necessary, for if a larger number of categories had been taken, there would have been too few individuals in some of the categories. Fisher\(^{(1)}\) actually recommends this method.

The only difficulty in carrying out the above sub-division is that in many cases a subject was given a zero (or normal) rating either by the teacher or by the tester. The whole 48 cases were thus distributed as follows -

<table>
<thead>
<tr>
<th>Blue Scores</th>
<th>(Testers’ rating)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ve</td>
</tr>
<tr>
<td>+ve</td>
<td>-</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>-ve</td>
<td>8</td>
</tr>
</tbody>
</table>

(1) Op. cit. P.
In order to convert this table into a 2 x 2 table the following procedure was carried out.

To begin with, there are 5 cases with + red scores but zero blue scores. We therefore divide these up equally, putting 2½ into the + blue score class and 2½ into the -ve blue score class. Similarly with the other cases where a pupil has one zero score.

Finally there are 8 cases who get double zeros. These were divided into 4 lots of 2, so that two cases went into the ++ class, two into the + - class, two into the -+ class and two into the -- class.

This will be made clear by the diagram below.

![Diagram](image)

Adding the contributions in each of the 4 cells we eventually get the following 2 x 2 table -
Now if there had been no connection between the two sets of ratings the numbers would have tended to proportionality. The expected numbers will be $a'$, $b'$, $c'$, $d'$,

\[
\begin{array}{c|c}
\text{ } & a' \\
\hline
b' & c' \\
\hline
\end{array}
\]

where $\frac{a'}{b'} = \frac{c'}{d'}$.

In fact $a' = \frac{25 \times 17}{48}$, $b' = \frac{25 \times 31}{48}$ etc.

The expected numbers are therefore

\[
\begin{array}{c|c|c}
8.854 & 16.146 & 25 \\
\hline
8.146 & 14.854 & 23 \\
\hline
17 & 31 & \sqrt{48}
\end{array}
\]

(The marginal totals remain as before).
As can be seen, the observed frequencies differ from the expected frequencies. To test whether the discrepancies are significant we calculate $\chi^2$. The easiest method of doing this is to find

\[
\frac{(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)}
\]

where $a$, $b$, $c$, $d$ are the observed frequencies. In this method it is not necessary to calculate the expected frequencies.

Putting $a = 5$, $b = 20$, $c = 12$, $d = 11$ we have

\[
= \frac{(5 \times 11 - 20 \times 12)^2 \times 48}{25 \times 23 \times 17 \times 31}
\]

\[
= \frac{185 \times 48}{575 \times 527}
\]

\[
= \frac{1642800}{305025}
\]

i.e. $\chi^2 = 5.42$

There is only 1 degree of freedom, so that from the tables (Fisher) we find that $P = .02$. That is, the probability of obtaining by chance a value of $\chi^2$ larger than 5.42 is only $\frac{1}{50}$. There is thus significance at the 5% level (the 5% point actually being only 3.84).

We have therefore demonstrated that there is a connection between the two ratings.

Conclusion.
Conclusion.

We did not expect to find an exceptionally high degree of association between the two ratings because of the variability of opinion likely between teacher and tester. We were therefore more than satisfied when we found significance at the 5% level.

There were 3 cases which showed a wide discrepancy between the teachers and the tester's ratings. All 3 subjects scored particularly low in the teacher's special subject, so that there was quite possibly a "halo" effect. Our calculation of course includes those 3 scores, and as we have seen, there is significance notwithstanding. It must be added, however, that the teachers' estimates were completed with obvious care. One pupil who does very well in school yet appeared to the tester to have a fairly low activity level. We found that the teacher also gave a low assessment on activity, even though the subject achieves a high class mark in her school subject. In this case, at any rate, the teacher avoided a "halo" effect in her estimate of the subject's general activity, as distinct from intellectual level, or achievement in her particular subject.
CHAPTER IX. SOME INDIVIDUAL RESULTS.

We shall now discuss a few of the individual profiles, in the light of our temperament assessment and their school achievement.

As regards the traits which we have picked out to form the profile, we should like to add that it is by no means intended to be a full description of each subject. It may not even be a "natural" classification, far less a good or true one, but it did seem to the writer to express to a fair degree her impression of the subjects. It is possible, after all, to divide a group of phenomena from very different points of view, and from amongst these different possible classifications, there is not just one true, and all the others untrue; they all are, however, more or less useful. Only judgments can be true or untrue. In a classification, however, one does not make a judgment, one merely classifies.

In general one has to beware of the narrowing of one's outlook, to which a definite classification, if one rigidly adheres to it, can lead. Reality, after all, is infinitely richer than any system, and it is of vital importance that we keep our eyes open to this richness.

Subject I./
Profile I

Average

SPEED Smoothness Flexibility Speed of Decision Manhiplation Freedom from Load Confidence Motor Imagination Care for Detail Sex Dominance

-3 -2 -1 0 1 2 3
Subject I.

This is the profile of a boy whose I.Q. is recorded as $130^+$.

Temperament:

Generally speaking this subject was one of the most poised in the whole group. Yet it was not the kind of poise which one would associate with a too early temperamental maturity. He used both hands throughout the whole battery of performance tests, the right hand being the predominant one. This would immediately tend to place him on the positive side for activity. Combined with this, however, was a smooth, level rate of movement, very slightly slowish, relative to the group in general. This led us to assess his activity as zero – adequate and well-balanced.

Along with this very well-adjusted level of activity, there was considerable evidence of flexibility, which enabled him to make an immediate adaptation to each sub-test of the battery. The moment one test was finished, he turned readily to the next, and gave it undivided and constant attention. It was this flexibility and freedom from load which lent his entire performance an air of lightness which redeemed it completely from being too adjusted, or too poised. He showed, also, a high degree of control over impulse, while never appearing in the least "inhibited"; what was manifested was merely the ability to stop and consider the consequences, before carrying out an action. He had a general air of confidence without the least sign of aggressiveness. Sex dominance brings this profile into proportion, we believe, for this boy had just enough delicacy in the handling of the material to give him the merest trace of femininity. His physique also was/
was supple and slightly slender; hair and skin texture were fine. It seems to the writer legitimate to include a slight description of the physical appearance in our report for it appears to give balance and perspective to our profile. We have thus assessed this boy as slightly below male average in sex dominance.

**Intellectual behaviour.**

Speed of decision was slightly below average. There was a very slight insistence on taking his own time to make up his mind. Coupled with great care for detail, this led to a highly competent performance, but being an able boy this lack of hurry did not in any way prevent him from being well within all time limits. In the formboard, for example, he placed the pieces smoothly and precisely. He made neither too many moves nor too few. He realised very quickly that some insets had to be reversed in order to go into the shapes. On this initial realisation followed learning by experience. At P.C.C. he carefully looked back to the diagrams before making his final placing, and made good use of the diagrams left before him at the constructional part of the same test. His performance with P.C.C. was one of the best the writer has seen so far. He learned it perfectly as he went along, and there was never a single false move. He gained the maximum score with Healy picture completion and gave good reasons for his choices. Kohs was done analytically. One could almost see him counting the number of cubes necessary; he always started at the top corners and worked downwards no matter what the design was. His total score on the performance battery was the highest for the group. Equally he gained a high score in the verbal group/
group tests 40, though not in the highest percentile. In Matrices he scored 57, (losing only 3 points) and in Space Perception 44. It is interesting to note, also, that his score in Matrices was the highest in the group. The test was given according to the original instructions of unlimited time. He actually was 8th longest in finishing, for knowing that he had ample time he paid full and careful attention to each item. (Incidentally we worked out the $r$ between speed and achievement with Matrices. It was $+.314$ and not significant.)

We thus concluded that this subject was temperamentally extremely well-balanced, and in an intelligent way. It was not just a case of being phlegmatically undisturbed, for his reactions gave evidence of very superior intellectual level as well as of motor balance. Also he appeared to have unusually good visual space perception, as was very plainly seen in P.C.C. Likewise, his verbal ability was on a high level, although not quite so high perhaps as the former. Imaginativeness and good reasoning were in evidence with Healy.

We thus interpreted his temperamental stability in terms of his mental capacity for the two were closely and very plainly well inter-balanced. At the time of the tests the writer had no idea whatsoever of I.Q. levels or of school achievement. It was arranged with the Rector of the school that, in order to prevent any "halo" effect the writer was to have no information at all about the pupils, until the testing was over.

This boy, as we have already noted, was recorded as having an
an I.Q. of 130+. He takes the first place in his section in school subjects, his general average being 89. He is very superior in all school subjects, the highest being mathematics and the lowest English (the latter of course was 80%). We might also have predicted this from his test scores! Temperamental stability was so much in evidence, that other abilities seemed to stand out all the more clearly in consequence.

As well as being praised for his good all-round scholastic achievement the writer was informed by one teacher that the main characteristic in his make-up was reliability. This was very much in accord with our own total impression and it looks as if here, at least, is one individual who makes adequate use of native ability counterbalanced by unusually favourable temperamental adjustment.

Subject II.

This is the profile of a girl, with the highest recorded I.Q. 135.

We have chosen to represent this subject, for she is an example of the discrepancy between I.Q. and performance test results which was one of our initial points of research. Apart from Healy (in which she scored $\frac{89}{97}$ and Kohn (in which she was average) she was below average in all the performance tests.

Temperament.
As regards activity, this subject was definitely on the positive side of the distribution. She used both hands liberally and went along at a speed definitely above average. She therefore gives herself a chance for careful thought, and care for detail was conspicuous by its absence. In the foreboard she joined an average score by sheer number of moves, as we have said this work.
Temperament:

As regards activity, this subject was definitely on the positive side of the distribution. She used both hands liberally and went along at a speed definitely above average. She was thus assessed as +2 on our combined activity variable. There was evidence not, however, of "controlled" speed, but of the general multi-movement characteristic of excitability. The actual moves were, however, fairly smooth. In confidence she was below average. This was seen in P.C.C. where, at step VII, she almost got the wood correctly placed, raising the bits in the direction of the correct position, but failed to push them down properly. Even with Healy, which she did well, this lack of confidence was in evidence, for she would have gained the maximum score if she had not replaced an already correctly chosen piece by another, over which she hesitated. In spite of her high speed, she manifested an above average degree of motor inhibition in patches. Taking it on the whole, however, we assessed her here as average. Generally speaking the subject gave evidence, temperamentally, of potential high activity, but lacking in directive force. It was of the dispersed, fluttering and excitable type which tends to lead to rather fruitless overt results.

Intellectual behaviour.

Subject's speed of decision in all tests (except perhaps Healy which was average) was far too great. She did not really give herself a chance for careful thought, and care for detail was conspicuous by its absence. In the formboard she gained an average score by sheer number of moves. As we have said this works/
works quite well, when the intelligence level is good, but tends to be rather painful to watch, with a subject of low I.Q. She made several mistakes with Cube Construction, through being in too great a hurry and simply not looking at the models carefully enough. In P.C.C. she was below average, but the writer had the impression that she could have done much better with care. In spite of her speediness, she lacked confidence in her own judgment, so that her fluency was more a matter of compulsive motor activity rather than assurance of her own capability. In Kohs she managed a slightly above average score, mainly because of speed and it looked too, as if she made a slight end spurt in the direction of carefulness, having observed that it was the last test. On the whole her visuo-spatial ability was definitely poor, or else her temperamental instability seriously hampered her performance. Such was our conclusion, at any rate. With Healy, however, she was most competent and the reasons given for her choice were intelligent and imaginative. Had it not been for an unfortunate second-thought choice, she would have gained the full score.

On the verbal group test she scored 44 (the third highest score) but in Matrices* 44 (below group average) and in Space Perception 39 (also below average). It was interesting, too, that she finished Matrices in 24 minutes (definitely quicker than the average of the group.)

* It is quite possible that for this subject Matrices was a mental performance test.
The writer concluded that this subject appeared to have rather poor visual space ability, but was of high level in "v". Temperamentally she was very excitable and this was brought out, it seemed, not only by the fundamental "performance" situation but because of spatial disability within the performance situation.

It was therefore most interesting to find out that this girl had the highest recorded I.Q. of the whole group, but that her position in class was 8th. Still more revealing was the information that her achievement in all school subjects was above average except in mathematics, in which she was 5 points below the class average. Again, we might almost have predicted this, from the test battery! Also a teacher informed the tester that this pupil was of the speedy, fluent type, showing great capacity to change readily from one thing to another. This too, was much in character with our own estimate.

More extensive investigation of this girl’s special abilities is indicated. There was, however, very distinct evidence of temperamental ill-balance, which may possibly be rather obscured in everyday life by high level of I.Q.

The Group in general.

The writer was really rather astonished by the wide varieties in temperamental reactions manifested with the selected group. All degrees of the traits of the two described subjects were in evidence. It was possible to pick out the deliberate and planful workers, from the impulsive and emotionally unstable types./
types. The latter was often manifested in irregular achievement throughout the tests, and in unit scatter even within a test. Signs of anxiety and insecurity were also noted, such as sighing and extreme tension in an unused hand. The verbal expression of such 'affect' might be spontaneous remarks such as "Oh! how stupid of me. I shouldn't have done that," and appeals to tester for an opinion on achievement.

In a general comparison with level of school achievement, the following broad temperamental divisions are represented in order of relationship with such achievement - poised, excitable, inhibited. The temperamentally poised subject utilises his mental ability to the full. The excitable subject is hindered by his excitability, but by force of accumulation of response, ranks higher than the inhibited type, who never seems to get anywhere near his potential level, and must therefore be consistently under-rated.

Comparison with Mentally Defective Group.

The writer thought that a comparison with a small M.D. group of the same age range would be interesting. She was able to test only 4 (10 had been aimed at, but a prolonged period of snow prevented more testing), but even with those subjects quite marked temperamental differences were in evidence with the performance test battery.

The profile represented is that of a defective girl of 13. I.Q. 75.

Subject I./
and only gained a score of 4 through some help. Kohs was like-impossibly difficult for her, only the first two steps being achieved and even those meagrely performed. She scored a little higher in Healy, because of the sense in which she was completing a picture. When she looked for a clock to fill one gap (the 8.30 clock being the correct choice) she could not get one of the clock pictures out of the selection box; it seemed to be slightly stuck. She therefore chose the alarm clock (the one required in the picture was obviously to be attached to the wall) as if it made no difference what kind of clock suited the picture. Her reasons for choices were very poor, and had actually to be rather pressed out of her.

There was no planfulness exhibited in this subject's performance, and absolutely no learning by experience. Also there was a kind of "mass" attitude to the whole performance battery, with no kind of differentiation as each test came along.

Finally the writer was most struck by the readiness of the subject to, as it were, give away some of her own personality to the tester. It seemed like a kind of self-extension, which seemed to be a function of low intelligence level, coupled with rather a "sweetly feminine" disposition.

She was described by her teacher as "an affectionate girl - inclined to be a little shy - capable at times of sustained effort - capable when looking after younger ones."

Subject II.

The second profile is that of a defective boy of 13.
I.Q. 67.
Temperament:

This is a more positive profile than the one just described. There was average speed of movement, but a very liberal use of both hands. There was much more confidence exhibited, than by the defective girl described, but it was tempered by a faint hesitancy, a lack of mental assurance. Motor inhibition was noticeable by its absence. This boy had an easily accessible temperament, although he frowned frequently when in difficulty.

Intellectual behaviour.

In nearly all the performance tests this subject gave the writer the impression of seeing to a certain extent what was required, but of being quite incapable of carrying it into execution. For example in Kohs, he appeared to see what was necessary and set out to carry out the thought, but unfortunately tended to forget which line or cube he was at, so that progress was slow and insignificant. In attempting to complete the diamond pattern in Kohs, for instance, he had no idea whatsoever of relative size or of counting the number of cubes required. He was certainly following out the pattern, but was quite incapable of achieving a "Gestalt." Although lower in I.Q. than the girl he had a better performance score because his visual spatial ability was better developed. He was by no means capable, however, of putting successfully into action what he seemed to "see", for intelligence was lacking which would direct his method.

This subject, too, was more childish by far, than any of the selected group of the same age. Generally speaking there was/
was a very definite slackness in movement, a lack of muscular co-ordination which differentiated both him and the girl from the selected group. Apart from the enormous difference in mental behaviour, there was this already mentioned air of faulty personality integration. Even the boy, who was a much more "positive" and fluent type than the girl, seemed to rely for his existence to some extent on the grown-up present.

The teacher's description of the boy is "Impulsive, yet a little shy - vigorous - inclined to be erratic - terribly "touchy" if opposed - goes home when he is checked. If he could control his temper, he should make a good skilled labourer."

In the light of this boy's activity level and interpreted relation to intelligence, we believe that the teacher's account tallies with his actual behaviour in the performance test situation.
We started out in our investigation with the idea of making an attempt to combine qualitative and quantitative methods in assessing some traits of temperament. Many of the purely qualitative surveys have given most important and valuable information in this field, but their value is mainly clinical; without some sort of objective standard against which they can be measured, the psychologist cannot safely generalise on the basis of qualitative information alone. But the psychometric procedure must not lead to psychologically meaningless abstractions. It is important that the rich complexity of human nature should not be split up and distorted into isolated statistical ciphers; and that broad quantitative generalisations should not be allowed to obscure the individual personality Gestalt.

First of all we made use of the questionnaire method in our preliminary experimentation, and concluded that only with reference to relatively 'non-affective' items, such as opinions and interests, do questionnaires elicit any sort of reliable information; and even with those, the subjects must be of good education. Unless utilised under personal interview conditions, questionnaires devised to obtain information of a more intimate nature, are likely only to measure what the individuals personally interpret as the meaning of the traits. The present writer is of the opinion that the questionnaire method should be scrapped in the meantime, except in so far as subjects are definitely asked to set down what they consider the meaning of such questions to /
to be, and not to what extent they think the traits are characteristic of themselves. In this way, we might possibly accumulate more evidence of certain "real" and fairly constant reactions.

As a result of watching the performance test behaviour of a girl suffering from high frequency deafness, the writer decided to use a battery of performance tests on a group of selected subjects to see if, on the basis of observation alone, and with as few a priori assumptions as possible, she could isolate some aspect of temperament while yet preserving its relation to the temperament "whole". In order, as it were, "to start with the sausage instead of the pig", she gave some performance tests and puzzles to a small group of highly intelligent, sophisticated adults to see what differences, if any, she could observe during their performance. Intellectually, their behaviour was all on the same high level, but different emotional reactions were revealed by means of the subjects' introspections, and above all there were well-marked differences in psycho-motor behaviour.

A battery of performance tests was chosen, difficult enough (theoretically) to reach adult level. A new performance test devised by the writer was included, fulfilling the criteria of both length and difficulty, and also designed to reveal (if possible) differences in 3-dimensional visual perception, in which point the investigator happened also to be interested.

In order to exercise some degree of control over intelligence (so that its influence in the temperament situation could to some extent be "known") the performance test battery was given to /
to a group of subjects who were 13 years of age, and not yet 14, and whose I.Q. level was superior. In spite of such homogeneity, very significant correlations were found between the tests used. The latter also included Progressive Matrices, N.I.I.P. Space Perception, and a battery of three verbal tests, these all being administered as group tests, and designed to give the writer supplementary information to the performance test results.

The 6 tests with the highest inter-correlations were subjected to a centroid factorial analysis and it was found that only the first factor loadings were significant; that is, the correlation matrix could be explained statistically in terms of one general factor. Psychologically this is meaningless. The method of Maximum Likelihood, invented to estimate the significance of factor loadings, might therefore be usefully applied to many previous factorial studies from which psychological conclusions have been deduced. These psychological assumptions have often been criticised, when founded on Thurstone's method of rotation, because of arbitrary and subjective psychological interpretation.

It seems to the present writer, that long before the validity of the psychology is questioned, the validity of the actual statistical method of rotation should be doubted, for if statistical results are not significant in the first place, no amount of alteration of dimension can make them so.

The complete test battery yielded information regarding intellectual behaviour in general, and special abilities in particular; and on the temperamental side well defined psychomotor individual differences.

The writer is of the opinion that much more intensive investigation of special abilities must be carried out, not only because of their educational significance in a vocational sense, but /
but because of the tendency of a large special disability to upset personality adjustment. A wide discrepancy between I.Q. and performance test score may indicate either special disability, or temperamental maladjustment (preoccupation with verbal symbols and unwillingness to carry thought into overt "performance" reaction, for example), or possibly a combination of both.

The whole field of imagery, which psychologists have tended to put on the historical shelf, should be re-opened and subjected to wide experimental investigation. Visual perception is likewise in need of further research, for the writer believes that variations on a general normal distribution scale of 3-dimensional visual perception are responsible at any rate for part of what we call spatial disability with visual material. One example of practical importance is the influence of such perception in the geometrical field of mathematics.

The writer erected a profile for each subject on the basis of general psycho-motor behaviour. She has done this merely tentatively to demonstrate what she believes is a possibility, arising from the performance test situation. She lays no claim to "goodness" or "trueness" of classification, but merely indicates what to her, at any rate, appears to be a useful description of her subjects. She does, however, put forward the suggestion that one very definite aspect of temperament can be tested. It is the general reaction which she designates "activity", both actual and potential, and has been assessed on the basis of mono- or duo-manipulation, combined with speed of movement. The writer is very well aware that she is laying herself open to the criticism of measuring only a single "level" of /
of personality here - the motor level. She is of the opinion, however, that this psycho-motor "Activity" is in this particular situation a generalized and fundamental reaction. Her reason for this belief is that in the performance test situation the use of hands is compulsory, only the particular mode of manipulation being spontaneous. The sole way in which the subject can put his thoughts into action with performance test material is by means of his hands, and since no time limit or need for speed is mentioned, he may proceed at his own rate.

The writer compared her own estimates on the general Activity trait with careful teachers' ratings, and using the chi-square technique she found that there was a significant relationship between the two. She believes that this assessment is of value for it permits a dynamic view of temperament. Activity by itself must in turn be interpreted in terms of I.Q. level, special abilities, and emotional range. The writer is strongly of the opinion that the mere estimate of I.Q. and school achievement at the 11+ educational stage, to take but one example, is a most inadequate sample of each child's personality. The sooner means become available for a much more extensive assessment, the brighter will be the outlook for adult adjustment.

The writer therefore puts forward the suggestion that a battery of performance tests is of value because of its 3-fold revelation of intelligence, special abilities and temperament. One great advantage lies in the element of control exercised over motivation, the uncertainty of the latter being the great stumbling block in so many investigations of temperament and general personality reactions. If the tests are difficult enough /
enough, they will demonstrate behaviour under conditions of
difficulty, and at the same time may reveal fundamental "constant"
tendencies, for the subject will usually assume that the aim of
the tests is to measure his achievement, and under conditions of
mental concentration may be expected to give some revelation
through psycho-motor channels of "real" generalized reactions.
The writer believes also, that in the performance test situation,
the essential unity of personality is preserved so that the
psychologist may make a total approach to the individual under
investigation.

Probably it would be doubly valuable to give each subject
two separate performance batteries at different times and take
results from the average of the two. The writer would also
like specially to mention the usefulness of the formboard type
of test, so long as it goes on long enough, as does the Kent-
Shakow. With this material the subject is at liberty to use
only one hand, if he wishes to do so, but there is ample scope
for the use of both, should such an inner temperamental
necessity arise.

In conclusion we must add that the method of assessment
we have described depends ultimately upon the comparability of
the behaviour shown in the miniature situation of the test with
that which takes place in the very much more complex social
setting of everyday life. The success of the method depends,
therefore, upon the possibility of being able to tap general
tendencies rather than specific reactions. With certain
reservations, the writer believes that the psycho-motor trait
Activity does give evidence of being just such a "real"
generalized tendency.
---|---
Cattell, R. B. (1936) | "Is the doctrine of instincts dead?" B.J.E.P. Nov. 1941.
Drever, J. (1921) | Instinct in Man. Cambridge Univ. Press.
Drever & Collins (1928) | Performance Tests of Intelligence.
Dellaert, R. | "L'Intelligence des anormaux du caractère". L'Année Psychol.
Garrett /


James, W. (1871) "Principles of Psychology. Vol. II".


Philp, H. L. (1936) "An Experimental Study of the Frustration of Will-Acts & Conation". Mon. XXI. B.J.P.
<table>
<thead>
<tr>
<th>Name</th>
<th>Year</th>
<th>Title</th>
<th>Publication Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piotrowski, Z.</td>
<td>1937</td>
<td>&quot;Objective Signs of Invalidity of Stanford-Binet Tests&quot;.</td>
<td>Psychiat. Quart., XI.</td>
</tr>
<tr>
<td>Pinard, J. W.</td>
<td>1932</td>
<td>&quot;Tests of Perseveration&quot;.</td>
<td>B.J.P. XXIII.</td>
</tr>
</tbody>
</table>