Psychology. — *Comparison of a few intellect-tests.* By Dr. D. WIERSMA. (Communication from the Laboratory of the State Asylum for Psychopaths, Leyden.) (Communicated by Prof. E. D. WIERSMA.)

(Communicated at the meeting of June 28, 1930).

The experimental investigation of the intellect may be set about in two ways, which differ from each other in principle. In the first place we can apply a method which gives us an insight into the intellectual performances of the test-person as a whole, which enables us to determine whether there is a debility or dementia, or on the other hand a more than normal endowment. What the nature of the defect is, in what direction a special talent lies, cannot, however, be determined with the aid of methods of this kind. The level of intelligence, determined by the BINET-SIMON method, tells us how far a child is in general ahead of or behind his age, but it does not give any clue as to the causes of the differences found. And the results obtained with tests like those of the American army inquiry, although well-suited for the very accurate demonstration of differences in intellectual development and endowment both within and without the normal limits, are restricted to a measure of the intelligence in general; they teach us nothing as to the separate elements which combine to form so complicated a function as the intellect. The importance of these general methods lies first and foremost in their practical application. For ascertaining whether a child should be at a school for mentally defective children, or in a class for specially gifted ones, the BINET-SIMON method is a greatly to be appreciated aid, and the brilliant results which psycho-technics have obtained with the American army inquiry and with allied series of tests, not only in the organisation of the army itself, but also in industry, hardly need to be recalled. In pathology, also, such level-methods can most certainly be applied with good results, as is clearly shown by the value which a test according to the BINET-SIMON method has for the diagnosis of debility or of dementia. Yet the determination of the level of intelligence in the case of adults is found to be nothing like so reliable as with children; it not infrequently occurs that there is a very great difference between the age-level at which all the tests still yield positive results, and that at which all the results have become negative. The calculation usually applied, according to which every properly answered question of an age level of which not all the tests can be calculated positively, counts for $^1/_5$ of a year, begins therefore to lose much of its value. I found, for instance, in the case of a patient of the age of 44, who was admitted into the Psychopathic Asylum, that all the questions for the age of 5 were properly answered, of those for 6 years

he answered 4 properly, of those for 7 years 3, of those for 8 years 2, for 9, 10, and 12 years each 1, whilst the tests for the age of 15 years were the first to yield nothing but negative results. According to the usual calculation we should therefore get an intelligence level corresponding to $5 + {}^{12}/_{5} = 7^{2}/_{5}$ years, but it need hardly be said that a result of this kind has little more than a conventional value. The question now is, to what such unsatisfactory results, which repeatedly occur in the case of adult test-persons, are to be ascribed. And then we have to remember that BINET and SIMON especially chose their tests so that as far as possible every kind of intellectual performances could be tested thereby: the repetition of dictated figures, counting backwards from 20 to 1, the giving of definitions, the criticizing of absurd sentences and of pictures with defects, the construction of sentences with three given words, are tests which have regard to very different sides of the intellect. And if a defective development or an acquired defect does not manifest itself on all these sides uniformly, then such discrepant results of the test are to be anticipated; the negative results have then to be obtained chiefly with similar tests. In the case of the patient mentioned it was more especially the tests which called for judgment and criticism in which he failed, such as the description of pictures, which usually give positive results at the age of 7, and the pointing-out of the difference between allied ideas and of defects in pictures belonging to the age of 8, whilst tests in which observation was an important factor: the arrangement of weights, which should be positive at the age of 10, and the offering of resistance to suggestion, when three pairs of lines are first shown, of which the right one is the longer, but afterwards three pairs which are of equal length, which should only be performed successfully at the age of 12, gave in his case positive results.

It must therefore be of importance to be able to test the separate functions also, which together form the parts of the intellect. And in this consists the second procedure which may be applied in the experimental investigation of the intellect, and which has the advantage of enabling us to determine which functions in the case of a feeble-minded or demented patient are chiefly deranged, in what direction an individual above the average is specially gifted. By testing each of the various elements of the intellect individually, and then combining the results obtained with one another, we are able by applying this method, just as well as by employing the first one, to form an idea of the intellect as a whole. But we can furthermore take into account the results of our tests for the measures to be taken in each case, which will be of especial importance if we wish to indicate the direction in which gifted young people can pre-eminently develop. Still greater is the theoretical importance of such an investigation. In the first place it is calculated to enable us to distinguish various forms of dementia from one another, since it gives an answer to the question whether in different forms the principal disturbance is not to be found in different functions. But besides this it is the only way in which we can

ascertain whether and to what extent particular elementary functions co-operate to produce intellectual performances. For this purpose it is not sufficient to devise a number of very heterogeneous tests; each of these must satisfy three demands:

1°. The activity demanded of the test-person must resemble that which we have per definitionem attached to a particular function.

2°. There must be a high degree of consistency between the results of tests which are directed to one and the same function.

3°. The consistency between the results of tests by means of which various functions are tested, must only be very slight; only in this case is a sharp distinction between the various elements of the intellect possible.

It is impossible within the scope of this paper to enter into the question whether it is justified and appropriate to reckon functions like the memory and the power of association to the elements of the intellect. Opinions differ on this point; the differences in opinion are partly due to a question of the interpretation of words, but they are also partly a question of principle. To decide on and defend a point of view would necessitate highly detailed expositions, which would here be out of place. I will therefore confine myself to discussing tests in connection with two functions, which are universally admitted to belong to the intellect: the *judgment* and the *imagination*.

For testing the judgment all kinds of different tests are in use. Miss BRUGSMA (3) made use of the definition method, ALBERTIJN (1) and ARKEMA (2) worked with pictures in which absurdities had to be pointed out. This latter method especially, which is easy of application, usually arouses the interest of the test-person to a high degree, and yields results which are pretty well independent of the subjective judgment of the tester, must be regarded as pre-eminently suited for the investigation of the judgment. ARKEMA easily found a number of absurd pictures, which, in accordance with BOBERTAG's requirements were correctly judged by 70 % of his normal test-persons of the working-class. At the psychopathic asylum, in addition to other tests in the same category, a series of 6 pictures is regularly used, in which the patients are required to point out the absurdity. This series consists of 1°. a cart drawn by three men, whilst a horse holds the reins, 2°. a fish in a birdcage, 3°. a hunter trying to catch a hare with his hands, although he has a gun over his shoulder, 4°. a clock with the long hand pointing to 12, but the small one between 1 and 2, 5°. a few trees lit up by the setting sun, one of which throws its shadow towards the sun, and 6°. a burning candle which throws a shadow of itself. It is found that the three last pictures are much harder to judge than the first two. Altogether they give results which agree very well with the other judgment tests. I have used them as material for comparison with the investigation to be discussed here.

For testing the imagination, also, various methods are in use. Prof.

WIERSMA (6), SCHULTE (5), and later ARKEMA (2) have made use of these methods with normal persons and with sufferers from various psychoses, and have found that they are very suitable for use as a method of testing the intellect. As is well-known, one procedure consists in showing the test-person each time a series of pictures in which the same object is represented in various degrees of distinctness, by focussing the camera with which they were photographed more or less sharply. The least distinct pictures are shown first, then follow pictures which gradually become clearer, in such a way that each series comprises six pictures. Note is taken of the number of pictures properly recognized by the test-person. With the other method the Dutch word "droom" (dream) is used, a note being made of the number of words with a meaning which the test-person, by omitting and transposing letters, can make in three minutes. The maximum number of words is 12, viz. moord (murder), oord (place), moor (moor), room (cream), oom (uncle), door (through), rood (red), oor (ear), drom (troop), dom (stupid), om (around), dor (arid), or 14, if such unusual words as do (musical note) and mor (present tense of the verb "morren" — to grudge) are counted, but normal test-persons of the working-class hardly ever reach this maximum; Professor WIERSMA found that they may be counted on to find 5 or 6 words. With medical students this number is, however, considerably higher; I found that amongst them the maximum was not infrequently attained. This difference must be attributed partly to selection, partly to better development of the function in question by practice.

Of late Professor WIERSMA has indicated still another method, which might serve as parallel test to the hazy photographs, by applying the principle of sensations constantly increasing in distinctness to the sense of touch instead of to that of sight. Generally known objects are wrapped for this purpose in flannel, in such a manner that the first object is surrounded by six layers, the second of the same sort by 5, the 3rd by 4, the 4th by 3, the 5th by 2, and the 6th only by one layer of flannel. The test person is given the packages in this order to feel with the 2nd and 3rd fingers of the right hand, and note is then taken of the number of objects that he can recognize properly. The checking of this test is the most important part of this investigation. What we had to do was to determine whether this test was really a test of the imagination, and whether there was a sufficient connection with the intellectual performances in general. With this object, after having experimented with 25 patients admitted to the psychopathic asylum, I calculated the correlation between the results of the test of the wrapped-up objects on the one hand and that of the hazy photographs, the words from "droom", and the absurd pictures each separately on the other hand. In order, however, to be sure that the three latter tests are really good comparative material, I also calculated their reciprocal correlations.

For although SCHULTE and ARKEMA had demonstrated the existence of a connection between the results of the photographs test, the "droom" test, and the general intellect-tests, they employed for this purpose the so-called

four-fields method. And although this is capable of showing an existing connection, it does not lend itself to a quantitive correlation calculation. I therefore prefer the formula of BRAVAIS-PEARSON. Since KAPTEIJN (4) showed that the fraction obtained by means of this formula is a measure of the relation between the number of common factors which have contributed to the results of two tests and the total number of factors which have been of influence, we know that in this correlation we possess a very accurate measure of the connection between the performances found. If all the factors which in the one test lead to a favourable result, do so in the other test also, and all counteracting influences likewise work in the same direction in both tests, without there being causes which make their activity felt in one test only, then the correlation between the two is $+1$, which of course in actual practice no more occurs than does the converse case, in which all the factors which bring about a favourable result in the one case, lead to an unfavourable result in the other, so that the correlation would be equivalent to -1. If the two tests are quite independent of each other, then the correlation is 0. If two tests are to serve for the investigation of the same function, then they must display a not too slight positive mutual correlation; this is usually required to amount to $+0.4$ to $+0.5$.

The tests with the wrapped-up objects were carried out with 25 test-persons, all patients of the psychopathic asylum. There were 7 series of 6 packages to be felt, in which were hidden: 1°. a matchbox, 2°. a key, 3°. a bottle, 4°. a button, 5°. a pair of scissors, 6°. a spoon, and 7°. a wire nail. The number of properly recognised objects amounted per test-person to an average of 28.5, distributed as follows over the various objects:

TABLE I.

Object	Properly recognized	Average per t.p.
Matchbox	135	5.4
Key	118	4.7
Bottle	112	4.5
Button	93	3.7
Scissors	88	3.5
Spoon	85	3.4
Wire-nail	82	3.3
Total	713	28.5

This result shows that the matchbox was the easiest to recognize, the wire-nail the hardest, and that the various objects in the above order —

which was purposely not adhered to during the tests with the various test-persons, but constantly changed — increased in difficulty. At the same time it also leads to the supposition that the test with the matchbox, in which more than 5 out of 6 packages were properly identified, is somewhat too easy for our purpose. This is shown still more clearly by the checking by the number of objects which should be recognized in each group by normal persons. For this purpose I only employed those 20 test-persons of whom it could with certainty be assumed that they had no disturbances of the intellect, and ascertained how many could in each case identify a particular phase. If a phase is recognized by at least 14 test-persons (70 %, that is), this performance must be regarded as characteristic for a normal person. We then find the following:

TABLE II.

Recognized with:	Matchbox	Key	Bottle	Button	Scissors	Spoon	Wire-nail
Flannel							
6 layers	**15** tp.—75%	9 tp.	13 tp.	7 tp.	3 tp.	3 tp.	5 tp.
5 „	17 „	**14** „ —70%	**14** „ —70%	8 „	5 „	8 „	11 „
4 „	18 „	16 „	16 „	8 „	9 „	10 „	11 „
3 „	19 „	19 „	17 „	11 „	**14** „ —70%	**16** „ —80%	11 „
2 „	19 „	20 „	20 „	**14** „ —70%	17 „	19 „	**15** „ —75%
1 layer	20 „	20 „	20 „	19 „	20 „	20 „	18 „
not recognized	——	——	——	1 tp.	——	——	2 tp.

It is here seen that the matchbox was recognized by 75 % of the test-persons even in its most difficult phase, which must be taken to indicate that this test is too easy. This objection does not apply to the other objects, only the bottle displays the peculiarity that the most indistinct phase is recognized by 13 test-persons, whilst the critical 70 % is only reached with the following phase. The lack of a sharp transition is a draw-back which renders this test also one of the least successful. The other objects however satisfy all demands that we are in the habit of making upon them.

It has now to be determined whether this experiment with the wrapped-up objects does really form an imagination-test. With this object I determined the correlation between the results of this test on the one hand and those with the hazy photographs, or the making of words from "droom" on the other hand. Of the hazy photographs 6 series of 6 pictures, representing 1°. a horse, 2°. a church, 3°. a mill, 4°. a dog, 5°. a bicycle,

and 6⁰. a ship were used. For the very reason that the correlation is a measure of the relation of the number of common factors to the total number that have been active in both tests, it is not necessary to exclude the disturbances of intellect from this determination. As it happens, the 24 test persons who, owing to fortuitous circumstances, were all that could be used for the correlation calculation, recognized on the average 19.4 of the 36 pictures, which agrees fairly well with the results of SCHULTE and of ARKEMA, both of whom found that a recognition of about half the pictures could be counted on. 23 test-persons were able to take part in the tests in which words were made from "droom": they found an average of 9.4 in 3 minutes, an exceptionally good performance, that is. I now found the following correlations:

Wrapped-up objects — hazy photographs: $r = +0.23$.
Wrapped-up objects — words from "droom": $r = +0.46$.

Although the latter correlation is satisfactory, we cannot rest content with this result, as there is an insufficient connection between the apparently parallel tests of the hazy photographs and of the wrapped-up objects, to permit of our assuming that they measure the same function. It is therefore necessary in the first place to ascertain whether both the photographs test and the "droom" test are to be regarded as imagination-tests. The performances of the test-persons in both tests meet the demands which can per definitionem be made. For the imagination is the function by means of which given ideational complexes are analysed into their elements and new combinations are built up out of these elements. When the word "droom" — apart from its meaning — is seen as a combination of letters, and from these letters other words with quite different meanings are formed, work is performed which agrees well with this definition. But the completion of an indistinct perception with various elements which make an intelligible whole with what has been given, as is done in the case of the test with the hazy photographs, is also an operation of the imagination in the sense just indicated. We have therefore in the second place to enquire whether a sufficient degree of connection exists between the results of the two tests. That this connection is not wanting has already been shown by the investigations of SCHULTE and of ARKEMA, but, as already stated, the extent of the correlation was not shown by their method. I therefore calculated the correlation according to BRAVAIS-PEARSON in the case of 23 test-persons:

Hazy photographs — words from "droom": $r = +0.08$.

This correlation is seen to be absolutely inadequate, from which we may conclude that only one of the two tests is to be regarded as a real imagination test, but that the other, although the imagination has some share in the result, is mainly influenced by other factors. In order, then, to determine which test is directed chiefly to the imagination, I made a comparison with the results of the judgment-tests which were obtained with the aid of the

above-mentioned absurd pictures, and calculated the correlation of the results of both tests with that of the absurd pictures, in this case, too, with 23 test-persons.

Hazy photographs — absurd pictures: $r = -0.02$.
Words from "droom" — absurd pictures: $r = +0.34$.

The first correlation is negative, but so low that we may safely assume that practically no connection exists between these two tests. The test of the absurd pictures — in which the critical faculty is the chief thing — meets very well the demands which may be made on a test of judgment; with the hazy photographs, on the other hand, the judgment apparently plays no part whatever. This latter experiment is thus the one that is to be regarded as purely a test of the imagination. It is quite a different matter with the making of words from "droom"; the fairly high correlation with the absurd pictures shows that in this test the judgment is a not unimportant factor, of even much greater significance than the imagination. Nor is this so very strange, for the alteration of the order of 5 letters undoubtedly demands a certain power of imagination, but is still a very simple operation, the combinations of letters so found must, however, be judged critically. The test-person has to ask himself whether they have any meaning and whether he has not already mentioned the word before, and he also has to take care that a letter which does not occur in the word "droom" may not have crept by accident into his words. We can therefore, by analysing more closely the performance required of the test-person, readily understand that in this test the judgment must be of importance.

If now we return to our wrapped-up objects, the question arises whether the fairly high correlation with the making of words from "droom" may not be due to the fact that the judgment is an important factor in the first test also. We can discover this by determining the correlation between the results of the tests with the wrapped-up objects and with the absurd pictures. This was also done with 23 test-persons.

Wrapped-up objects — absurd pictures: $r = +0.26$.

We see that this correlation is, it is true, not sufficient to justify our regarding the object-test as a test of judgment, but that it nevertheless has a certain significance. It is even seen that judgment and imagination are in this test of about equal importance. At first sight this may seem surprising, since there seems to be an absolute parallelism between the hazy photographs which give the test-person ever-distincter visual perceptions, and the wrapped-up objects, with which it is tactile perceptions which constantly become clearer, but I believe that the fact is to be explained. Human beings are wont to live principally in visual perceptions. They have in this way learned to interpret their visual impressions immediately. With tactile impressions this is not the case. A perception comes about by our arranging the separate sensations and combining them into a particular image. From

childhood we have practised this, so far as our visual sensations are concerned, so that we immediately recognize a certain form even in indistinct combinations, and have no need to ask ourselves consciously what we really see. On the other hand such an indistinct image will require certainly some amplification, for which we have to call in the aid of our imagination. In our tactile sensations, however, we are nothing like so much at home; if they are indistinct, we shall certainly, owing to lack of practice in this direction, have to make great efforts to try to realize what we really feel and exactly what shape corresponds to our impressions. In addition to this, we are able, as it were, to take in at one glance a picture that is not too large, whereas when feeling a package we get a number of consecutive impressions which combined have to form our perception. We therefore have first to judge our touch-impressions, as accurately as possible to determine to what shape they correspond, before the second phase, the amplification of the imperfect image by means of the imagination, can begin. In view of this, therefore, it is clear that in the test with the wrapped-up objects the activity of the imagination is preceded by an activity of the judgment, which is of equal importance for the bringing about of the final result. With the test of the hazy photographs, however, this activity of the judgment has become so easy owing to the tremendous amount of practice which we have all had in the course of our life with respect to our visual perceptions, the judgment comes about so rapidly and so very unconsciously, that this function is of no more importance for the result of the test. The difference between the two tests is, I think, sufficiently accounted for in this way.

With this, then, the place of the test with wrapped-up objects amongst our other intellect-tests is indicated. They do not form a test with which one particular element of the intellect can be investigated, but they certainly have a value as an intellect-test in general. The functions of judgment and of imagination are of about equal importance for the result. Of the separate objects which were used, the matchbox and the bottle were found to be the least suitable, all the others, however, were very suitable; the key should be identified when wrapped in five layers of flannel, the scissors and the spoon when packed in three layers, the button and the wirenail, finally, when surrounded by two layers.

The tests of the hazy photographs and of the absurd pictures proved to have practically no common factors; the former is to be regarded as a pure imagination-test, the latter as a measure for the judgment.

Finally, the making of as many words as possible from "droom" proved to demand an imagination-element, but to be dependent to a much greater extent on a well-developed judgment.

LITERATURE.

1. ALBERTIJN. Bijdrage tot die methodiek van die intellektsbepaling. Diss. Groningen 1917.

2. ARKEMA. Experimenteele psychologische onderzoekingen bij normalen en geesteszieken. Diss. Groningen 1921.

3. Miss BRUGSMA. De bruikbaarheid der definitiemethode in de kliniek. Diss. Groningen 1921.

4. KAPTEYN. Definition of the Correlation-coefficient. Monthly Notices of the Royal Astron. Society 72 (6) p. 518. Quoted by:

HEYMANS und BRUGMANS. Intelligenzprüfungen mit Studierenden. Ztschr. f. d. angew. Psychol Bd. 7, 1913, S. 317.

5. SCHULTE. Onderzoek over phantasie en intellect bij normalen en geesteszieken. Diss. Groningen 1918.

6. E. D. WIERSMA. Intelligenzprüfungen nach BINET und SIMON und ein Versuch zur Auffindung neuer Tests. Ztschr. f. d. angew. Psychol. Bd. 8, 1914, S. 268.