PICTORIAL COMPLETION TEST II¹

By WILLIAM HEALY, Judge Baker Foundation, Boston

INTRODUCTION

It seems clear that any fair attempt to measure intelligence must include well-directed estimations of apperceptive abilities. Surely the capacity for "putting two-and-two together" in the realm of thought, the ability to "turn things over in the mind," to "rationalize" about perceptual material with what may be drawn on from the ideational and other memory stores of the mind is of the very essence of intelligence and of very great importance in civilization. This, if anything, may be spoken of as a higher mental power.

My interest in a pictorial completion test dates from appreciation of the probable great value of testing apperceptions and from the realization that the language completion test (as used by Ebbinghaus and developed since by others) involves several features which interfere greatly with its early intended use as a test for apperceptive abilities. First, there is the outstanding fact of the immense differences in the opportunities of individuals for acquiring facility with the use of our language, not only differences in formal educational advantages, in family and other cultural opportunities, but also, here in America, in the extent to which English is actually spoken at home or—as unfortunately with some of our immigrant peoples wedded to their own institutions—even in school.

¹ This test was exhibited first at the December, 1917, meeting of the American Psychological Association. Soon afterwards the War Department requested the use of it, and, of course, further publication at that time concerning the test became undesirable, even of the scoring values and norms which were at once worked up for the purposes of the Army psychological examinations when the test became part of them. Then the illustration was copyrighted in its several parts in 1918 in attempt to prevent the unfortunate newspaper publicity which befell my first Completion Test. The Judge Baker Foundation (40 Court St., Boston) has recently undertaken to have the test conveniently mounted and boxed and obtainable for a reasonable price.

(Since this article, including the above foot-note, was written, I have observed that this test, including the small plate of the whole series of pictures, which we have purposely refrained from publishing ourselves, and a copy of the scoring values calculated by us, has been published without any mention of authorship in a book entitled, "Army Mental Tests," by Yoakum and Yerkes; Holt & Company.)

On account of this total situation it is impossible to utilize the language completion test as an apperception test equally fair for all.

And whenever the verbal passage becomes really difficult, as in the harder parts of language completion tests, unusual words and those with special shades of meaning increase the difficulties for interpretation of the test results, throwing the advantage strongly to the individual who has had specially good training in language.

Second, I stated my conviction long since, giving instances bearing on the point, that language ability was to be regarded as a specialized ability, sometimes very poorly correlated with the findings on other intelligence tests. We now have considerable data bearing on this. Just as one may possess to a superior or very inferior degree the power of musical expression, so language may "flow trippingly" from the mind or one may by nature halt with words. In the direction of either extreme there is vitiation of the language completion test as indicative of apperceptive ability in general.

Third, the number of details necessary for a desirably complex situation for testing apperceptions can only be presented to a reader at the expense of taking considerable time for the reading and, through this, there is involved another well-known variable, the memory powers or memory span, in this case the memory span for ideas, as they are represented one after another in a long passage.

It suggested itself, then, that a picture completion test might be much fairer as a test for apperceptions, eliminating to a large degree the factors of training and of special ability or disability in language which, as a mode of expression, could be only incidental to the main purpose of the language apperception test. And a well devised pictorial test certainly would give opportunity for presentation of a wealth and exactness of ideational detail at a glance which would require pages of text adequately to cover; as complex a situation as desirable could be depicted without any strain on the memory or any troubled traveling of the attention back and forth, as over the printed page.

In a picture we could present ideas of a simplicity comparable to those of a simple short sentence, or one could illustrate situations too complex for any ordinary rapid solution. And then, as I demonstrate in Pictorial Test II, by the use of the serial pictures one can carry along ideas as they were carried along in the earlier language completion tests where words were elided from a narrative of some length.

Ideas from which to choose in filling the spaces of incompleting sense are represented by a stock of small pictures, and, as one selects verbal ideas from his mental stock in the language completion test, one chooses pictures to fill in what has been cut out of the main illustration.

DEVELOPMENT OF THE TEST

The first pictorial completion test which I devised some years ago² soon became widely used and commented on, particularly in the monograph devoted to it by Pintner and Anderson.³ That it seemed to others to embody a sound idea for a good mental test, and that this first picture showed itself to be crude in several features, both for interpretation and scoring, stimulated me to develop plans for a much better test, of the same sort, one that should involve both easy and decidedly difficult problems for apperception. This appeared to be possible by thinking out a pictorial narrative, series of pictures with elements for interpretation carried over from one picture to another just as elements of significance for understanding a narrative are carried over from a sentence or paragraph to later passages, easy or more difficult.

The fact that this idea for the test has not been carried out by others is very likely due to the difficulties and the amount of effort involved.⁴ For the development of a series

² Described in *The Psychological Review*, May, 1914

³ The Picture Completion Test, by Rudolf Pintner and Margaret M. Anderson, pp. 101; Warwick and York, Baltimore; 1917.

⁴ All told, a very considerable task has been involved in the development of this second pictorial test, an effort quite beyond the ordinary capacities of any one or two persons. It has been through the fine cooperation of a group, the members of which have given without stint of their time, that the work has been done. Dr Bronner has been foremost in advice and effort at every stage of the practical development of the test. Indeed, three of the eleven parts of the picture are entirely her suggestions, embodying better ideas than I had been able to advance. The illustrator, Miss Marguerite Davis, brought to the task unflagging zeal as well as artistic skill. The important mathematical suggestions and computations for scoring values were made by Mr. Herbert Sturgis. The entire staff of the Judge Baker Foundation gave itself over to obtaining results on the test both on many trial pictures and on the test as completed, for the establishment of norms. And groups of various nationalities and ages and grades of educational advantages were tested. Much work was undertaken in the interests of the army psychological service during the early part of 1918 in order to establish values for scoring and norms of performance for various groups on this test. Very valuable service in this was rendered by Margaret Fitz and May Bere. A number of outsiders also aided at different stages, a notable amount of work being done by Miss Katherine Coveney.

PLATE I

Sketch from the colored illustration of the simplest picture in the series, the demonstration picture, with the inserts which particularly relate to this picture. Of course only one is quite correct, but the others can be inserted with some show of reasonableness.

of satisfactory ideas a considerable task is demanded of the psychologist, preceding combined work with the artist. One of the main practical obstacles is found in the necessity for devising the correct proportionate size and an insignificant background for all the elided parts, so that there shall be no fitting in of the pieces according to the form of the piece, or by joining together parts of the picture. So far as the size of the pictured object for placement and the immediate background is concerned, any one of the related stock of idea-objects must seem reasonably placeable in any one of the spaces to be filled. Then, too, the carrying over of a central idea from picture to picture with broad illustrative conceptions, but with an exactness of representation that shall satisfy critical interpretation is no easy matter.

DESCRIPTION

Pictorial Completion Test II consists of an illustration to be divided for convenience into two or three parts when mounted. It is composed of a series of eleven pictures $5 \times 3\frac{1}{2}$ inches each (one, very simple in idea, merely for demonstration purposes), representing in sequence situations or events occurring during a day in the life of the boy who is depicted in each picture. From each picture there is a piece cut out, one inch square; the picture to be completed by selection of a square from the accompanying sixty small illustrations, only one of which represents the object which is needed to exactly fulfill the idea of the given picture. In every instance several of the small pieces depict objects designedly logical in part according to whether various ideas suggested by the given picture are taken into account, but not wholly correct, and hence may partially comply with the requirements for thinking out the situation as pictured. A few of the inserts are purposively suggestive. Of course, most of the sixty small pieces are quite inconsistent with the logical needs of any given one of the eleven pictures.

The ten test pictures present situations planned to be of greatly differing difficulties of solution, and they have worked out so well that children of eight years generally get a credit score, while many adults, even of very superior achievement, find it difficult to achieve a perfect score.

The drawing of the illustrations is satisfactory and the bright coloring stands out well as done by the three-color process. The picture is varnished and should be mounted on three-ply scroll-saw wood. The sixty pieces are to be numbered on the back for scoring purposes and are given standard

placings for presentation, according to numbers to be placed on the floor of the box in which they are kept. The pieces which with some reason might belong in any given picture have a certain amount of grouping in our arrangement so that the search for related "ideas" may not be unduly difficult.

(For those who get the illustration not cut up it should be stated that the numbering of the 60 pieces proceeds from the left top across the page and is repeated for each one of the six lines of 10 pictures. Thus, No. 1 is the scissors, No. 10 the picture and frame, No. 60 the top. Of the larger pictures, the boy getting up in the morning is the demonstration piece; No. 1 is the breakfast scene; No. 2, on the way to school; No. 3, at the school house steps; No. 4, the dressing room; No. 5, the schoolroom; No. 6, boys looking up at the sky; No. 7, the street scene with the policeman; No. 8, returning from the baseball game; No. 9, two boys on the side-walk; No. 10, evening scene in the home.)

DIRECTIONS

The entire picture, parts 1 to 5 forming the top row and parts 6 to 10 below them, should be placed in a good light before the subject. The sixty small pieces in the standardized arrangement are presented in the most convenient place where the subject may see them well.

The procedure is then very simple, but should be according to rule. For very dull persons repetition and increased emphasis of the following verbal directions may be needed.

"Here is a picture—it begins here" (pointing to demonstration picture) "where the boy is getting dressed. It shows the **same boy**, remember, the **same boy**, doing one thing after another during the same day." (Point along the first row, then along the second, to indicate clearly the sequence in which the pictures come.) "You see in each picture a piece is missing. Here" (pointing to them) "are a lot of small pieces, they fit in any of the spaces. But there are more pieces than you can use. The point is to pick out the piece that you think is needed, that is best, to complete the sense of the picture. For instance, What is gone here?" (pointing to demonstration picture). "Yes, a shoe." If incorrect answer is given, which is very rare, Examiner says, "No, he is dressing and he is stooping for his other shoe." "Now, which is the shoe that he must have?" If correct shoe is selected

Examiner says "Yes. This one" (pointing to low shoe) "wouldn't be right because he must have a high shoe to match the other one."

"Now, that is the way each is to be done. There is always some piece that is the very best one; you can tell which it is by studying the picture. Now go ahead."

The main point is to give no help after the first explanation. Nothing further is said while pieces are being inserted except that examinee, if he asks, is told that he may change pieces. After all empty spaces have been filled examiner asks if examinee is quite satisfied, if not, he is told that he may make further changes. When examinee indicates that he has finished as well as he can, time is recorded.

Only the final placings are scored, but record may be kept, if desired, of the changes. Next to the numbers of the pictures in order may be recorded the number or name of the piece inserted.

Twenty minutes is maximum time allowed.⁵

The total score on the test is the sum of the values of the pieces as finally placed. If all are correct the score is 100.

SCORING VALUES

The value of minus 5 is to be given to all placings where in the table below no numbers are inserted. These placings represent marked absurdities. The numbers in bold type are for correct placings.

Questions will determine in suspected cases whether a correct placing has been accidental. It happens rarely, as when the correct window in IV is placed, although "any window will do." The score of that placing, obviously, should be no more than the verbal judgment calls for, e.g., the value of an incorrect window in IV.

⁵ In the army examinations only 10 minutes was allowed on account of necessity for saving time. This was rather unfortunate, although, as a matter of fact, it is decidedly rare that longer than this time is desired by the subject. However, the differences that there may be is shown by the fact that one of the most brilliant men in America took the full 20 minutes and wanted more in which to finish Picture 10, finally leaving a piece in place that gave a total score of 90—the relation of the sewing to the meaning of the picture he had not yet apperceived.

TABLE I

VALUES OF PIECES AS PLACED IN PICTURES I TO X

Pieces	I	II	III	IV	V	VI	VII	VIII	IX	X
1.....					2					12.5
2.....	0	0	1	2		0	0	0	0	0
3.....							1	15.5		
4.....					0					6
5.....							0	0	0	
6.....	0			0						
7.....	0	0	1	2		0	0	0	0	0
8.....	6			2						
9.....		5	0		0					0
10.....	0			1						
11.....	1			8						
12.....		0			0					1
13.....		5	0		3					1
14.....							1	6		
15.....			1			1	0	0	0	
16.....							0	0	0	
17.....							1	6		
18.....			0			9.5		0		
19.....		2			0					0
20.....										
21.....		0			1					0
22.....					2					6
23.....	1			18						
24.....					0					6
25.....							2	0	0	
26.....			0			4	0	0	0	
27.....		5	0		1					0
28.....			0			4	0	0	0	
29.....										0
30.....	2			2						
31.....	1			8						
32.....		0			7					0
33.....							1	7		
34.....							5.5	0	0	
35.....			11		0	0		0	0	0
36.....			1			2	0	0	0	
37.....		0	4		0	0				0
38.....										0
39.....							0	0	1	
40.....										1
41.....		0			2					0
42.....	0			2						
43.....		0			2					0
44.....			1			1	0	0	0	
45.....	3			2						
46.....	0	0	6	0	0	0				0
47.....		5	0		1					2
48.....							0	0	5	
49.....		10	0		1					2
.....	0						0	0	1	

TABLE I—Continued

Pieces	I	II	III	IV	V	VI	VII	VIII	IX	X
51....	0	0	1	2	.	0	0	0	0	0
52....	1	.	.	8	.	.	.	.	.	.
53....	0	0	1	2	.	0	0	0	0	0
54....	.	.	.	.	.	0	0	0	0	.
55....	.	.	6	.	0	0	.	.	.	0
56....	.	.	.	.	.	.	2	0	0	.
57....	.	.	.	.	.	0	0	0	1	.
58....	0	0	1	2	.	0	0	0	0	0
59....	.	.	.	.	0	0	0	0	2	.
60....	.	0	2	.	0	0	.	.	.	0

The above scoring values were computed from the results obtained by testing 1,542 individuals consecutively, all of them presumably of normal mentality. We did not include records of the feeble-minded or psychotic.⁶

The task of working out proportionate scoring values of correct placements upon the basis of all pieces accurately placed, giving a score of 100, was assigned to Mr. Herbert Sturges. He worked first with a series of 257 records and then with 1,542, including the first 257. The resultant values differed very slightly for the two series, indeed, five of the correct pieces remained practically unchanged, less than .5 each, the greatest alteration being from 12.35 to 11.24 as the value of the correct placing in picture 3.

The method applied to the problem was that of weighting each correct solution inversely to the frequency with which it was correctly solved. The working out of the method is as follows: Opposite each frequency write its reciprocals; find the percentage that each reciprocal is of the sum of the reciprocals.

The rule stated is: Value *a* for problem A is the percentage that the reciprocal of the frequency of correct solution of A is of the sum of the reciprocals of all the frequencies of correct solutions.

⁶ Here I wish to acknowledge the value of Pintner and Anderson's criticism concerning the inadequacy of the scoring methods and norms for my first completion test. In turn, it was this criticism that challenged Dr. Bronner and me to attempt to develop a less clumsy system of scoring than the one devised by these authors. The main point they make is thoroughly valid; many tests do need standardizing for age and other groups. But on all except group tests the necessary labor—over a thousand hours were devoted to standardizing this our present test—and the requirement for much coöperation stands in the way.

TABLE II

VALUES FOR A SERIES OF 1542 RECORDS

Picture	Correct Placings	Percentage of total correct placings	Reciprocal of frequency	Score	P. E. of Score
1	1130	14 44	0 000885	6. (5.77)	0.06
2	655	8 37	0.001527	10 (9 96)	0.20
3	580	7 41	0.001724	11. (11 24)	0 25
4	363	4 64	0 002755	18 (17 96)	0.56
5	925	11.82	0 001086	7. (7.08)	0 10
6	683	8 73	0.001464	9 5 (9 55)	0 18
7	1228	15 69	0.000814	5 5 (5 31)	0 05
8	416	5 32	0 002404	15 5 (15 67)	0.46
9	1325	16 93	0 000755	5 (4 92)	0 03
10	520	6 65	0 001923	12 5 (12.54)	0 30
Totals	7825	100 00	0.015337	100 00	

The question of giving values to solutions other than the one correct solution of each picture could not be answered by mathematical considerations based on the frequency of such solutions. Arbitrary determinations embodying as much common sense as possible were invoked. No approach to correctness, it was felt, could be worth more than half as much as the correct solution. So the pieces which were planned to represent ideas similar to the single correct piece but missing the real point were given values of half or a little less than half. Thus the value for correct placing, the red book, in Picture 2 is 10 (the nearest half decimal to 9.96). But if one finally chooses the blue book or the two books or the lunch box or the box of pencils the score is 5. Of course it could be nothing else than the red book if one apperceives the situation, the boy having dropped something out of his strap, and his books, plainly colored, being shown strapped together in the previous picture. The only other possibility is a footprint on the ground, and that receives a score of 2. In the other pictures the commonsense of the situation and the possibilities differ and various arbitrary rulings concerning scores have been made accordingly—always based, however, on the original computations of correct scoring values.

If a piece of indifferent meaning, such as a blank in Picture 2, is left in place, or if, as extremely rarely, the space is left unfilled it has been decided that the score should be 0. But a large share of the possible placings are obvious absurdities and logically should be penalized. It is pretty clear just which these absurdities are. It was decided that minus 5 should be the standard score for these absurdities.

Though the scoring of these latter classes of placings is thus numerically arbitrary, still it must be remembered that the scoring as a whole is based upon mathematical computations of fair values according to ascertained relative differences in performance and that once this was done, the minor values as established are as fair for one person as another. (Of course I realize the suggestion value or the association value of any piece is never the same for one person as for another, but the same might be said to a greater or lesser degree about almost any mental test.)

TABLE III
PERCENTILES FROM THE ABOVE CHART

Ages	7	8	9	10	11	12	13	14	15	16	17-20	20-50
75 percentile	24	41	48	59	63	66	69	72	76	75	76	78
Median	9	27	37	47	54	55	58	62	64	66	65	65
25 percentile	-6	7	23	32	41	45	50	52	54	54	54	54

Special educational advantages. Whatever the findings for those who have had educational advantages above high school it would hardly be safe to allege that any good score would be due to such advantages rather than to native ability. It is more than conceivable that a considerable number of those who go further with their schooling do so because of their already demonstrated ability. Nevertheless it was felt to be much wiser to keep separately the 231 records of this group.

TABLE IV
Scores Obtained No. of Subjects

91-100	57
81-90	54
71-80	47
61-70	42
51-60	21
41-50	8
31-40	2
	<hr/>
	231

75 percentile = score of 90
Median = score of 80
25 percentile = score of 68

It is interesting to compare these much higher percentiles with those of the main group. And then it should be stated that a perfect score was obtained by 25 (11 per cent) in this last series as against only 4 (1.2 per cent) of 320 of the same age-group (above 17 years) in the main groups.

NORMS OF PERFORMANCE

Age-groups

FIGURE 1

Age distribution of total scores, 160 for each age-group (records of the feebleminded and of those with educational advantages above high school are not included).

While complete significance cannot, of course, be given to the findings for age-periods with the higher educated group on account of uneven and too few numbers, yet it is striking that 20 per cent of 65 individuals 35-50 years of age gave perfect scores and that their median score was 85, while 7.3 per cent of 164 with ages 19-35 did perfectly, with a median score of 76. It would seem as if among well educated people youth did not have the best of it in this test.

The Mentally Defective. Records of the performance of the feeble-minded on any one test show the futility of attempting to establish for any single test norms which are characteristic of the mentally defective. Mental defectives differ greatly among themselves in their special capacities and in other important features of mental life than are expressed by general age-level ratings on a given "intelligence scale." Some are feeble-minded because of certain defective memory powers, others with good memories have poor reasoning ability, some show great unevenness in their mental powers, some are below par evenly in their various mental capacities. When really good chapters on analytic psychology of mental defectives shall have been written, we may learn something of the complexity that most certainly exists in the mental life of the feeble-minded, which to many seems a simple affair. As it stands, we should scan the records of defectives in general on any special test with the knowledge that they represent the performances of individuals with age-level ratings differing widely from actual age experiences and from varying world experiences. Or the records may come from those who are not clearly feeble-minded except insofar as social behavior is concerned (institutions for defectives sometimes retain such cases) or from schizophrenics who are not infrequently held among the feeble-minded.

Even on a test involving such complex mental functioning as the one under consideration we may find great differences among defectives, at least of the higher grade, who have equal rank on age-level scales. However, the ordinary observation that some show clear mental grasp (apperception) of their own disabilities and various other practical situations, and other defectives of equal age-level rating by the Binet scale or its modifications fail thus to size up the real situation in regard to themselves and other affairs, ought to make the fact of their mental difference in this particular altogether plain and suggest at the outset that their scores on a special test for apperception might differ widely.

Other Groups. From records obtained by way of experiment, but too few for safe deductions, we have noted several suggestions for special studies. Persons who have seen very little of books and are *unaccustomed to observe pictures* understandingly or very little in any way, all of them being, with us, illiterate foreigners, were at a great disadvantage and clearly could not interpret the sky or the interior of a room as such, to say nothing of finer points in the pictures. *Psychotic* individuals have made all sorts of records, as may be supposed—any study of their performance would have to be by types and stages of psychoses. The apperceptive ability of *old age* is a matter of special interest and perhaps might turn out to be of great practical importance, because of the value that is given by tradition and social usage to the opinion of men of age, especially as they function on the bench or in administrative positions.

DISCUSSION

The universal interest evoked by the test is readily understandable. As a series of incomplete pictures it challenges the performance of everyone, for even the mentally defective, above low grades, see enough of the significance of what is depicted so that they become interested and are able to score as having observed correctly and reasoned correctly about some of the material in the pictures. An opportunity like that of everyday life is afforded them for some sort of not altogether unreasonable performance. In this, namely, in developing a test that in given situations for apperception should have simpler elements capable of being readily rationalized by very immature minds and at the same time embodying features of much greater difficulty, up to those relationships which require close thought by competent reasoners, one of the foremost aims of this test has been realized. And it is a fact that even the cleverest, when they note the possibilities of interpretation, enjoy working out the problems of the test as tasks calling for good effort.⁷

The ideas as depicted in the pictures finally drawn are so clear that they present no logically equivocal issues. This has been proved for us not only by the records of performance, but also by the comments of those doing the tests and their acknowledgments as they have later seen the point for correct solution; even after having themselves earlier failed to apper-

⁷ We have just heard that the work with French children on this test as done by Marion Pilpel through Simon's clinic in Paris has been very satisfactory. The situations have been readily grasped by these children.

serve some essentials. Indeed, it is interesting to note that in the two pictures awarded the highest values because of the infrequency of their correct solutions, the correct insert is indicated with more exactitude by objects clearly shown in the main pictures than is the case in any of the other pictures. There is less chance of equivocation in these two pictures because the possibilities of solution, once grasped, are so immediately related to the factual necessities of the situation. The most common failures, then, are directly perceptual and inferential in character, involving faulty observation or faulty reasoning concerning objects and phenomena universally known and among the most plainly observable in the illustrations.

In even the simplest analysis of the test, as a test measuring something much more than primary mental powers, such as memory, we encounter the facts, then, of its easily observable and readily recognizable material, of the facile perceptibility of the elements in and for themselves. The difficulty of the problem comes in rationalizing the value of each whole situation, of the worth of the various elements and of the cut-out objects (the store of ideas) as possible solutions in the given situation—and this is apperception.

Through its presentation of problems that are visual, non-language, ideational, pictorial, a representation of everyday scenes and activities, this test brings into play an important mental ability—whether specialized as being related to observational and pictorial material or capable of wider interpretation—an ability that we would eagerly know more about for many practical purposes. The reaction expressed by one well-known psychologist on first seeing the test, "Well, if I were hiring people to assist me, I think I should like this test to hire them by," may be paralleled by many another opinion of like import from those who have been using this among other tests. Noting that the higher records on the test, we may take the perfect score, 100, for example, have been done by only very occasional youngsters of high ability, as proved by other tests, and by some adults who have established reputations for achievement, and these records have been equalled very rarely indeed by others, for instance, by a domestic cook, we may look further and generally discover in these cases (as in the latter instance where this cook was found to be well known as having exceptionally good common sense and judgment) that there are obtainable evidences of the correlation of such a performance with otherwise known good mental capacities. The powers that this test in some measure gauges are certainly those allied with better human achievement.