THE MEASUREMENT OF DEMENTIA *

By M. B. Brody, M.D., D.P.M.,
Senior Resident Physician, Runwell Hospital, Wickford, Essex.

[Received December 20, 1941.]

Hitherto two methods have been commonly used to measure dementia—by means of numerical measures of scatter, and by means of the discrepancy between scores on vocabulary and other tests. In the present investigation the value of the most popular techniques with both methods was examined, and a new short measure of dementia was evolved.

Subjects taking the Stanford Binet and similar tests rarely pass abruptly from complete success at and below a certain M.A. level to complete failure at the next and subsequent levels. For example, a subject with a total M.A. of eleven years may succeed in all the subtests at and below the VIII year level and fail all the subtests above the XIV year level, while between these two levels he succeeds in some and fails others. The range of partial success and failure is called the " scatter." Many psychometrists believe that unusually wide scatter indicates psychopathy or intellectual deterioration. The various methods for measuring scatter are conveniently summarized by Harris and Shakow (1937).

A long-standing observation on the results of intelligence tests in dementia is that vocabulary ability remains comparatively unimpaired until a late stage when the subject is unable to solve other tests. Therefore, it has been claimed, vocabulary score may be taken as a measure of the subject's pre-morbid intelligence, and the discrepancy between his vocabulary ability and ability on other tests as a reflection of the difference between his former and present intelligence, thus providing a measure of his dementia. The Babcock test is the best known of several techniques, based on this principle, for the measurement of dementia.

Subjects and Tests.

Subjects.—The subjects were 83 mental hospital patients divided into four groups (Table I). Those in group A were not demented, those in groups B, C, and D, " mildly," " moderately " and " severely " demented. No test showed significant differences between groups B and C, which were therefore combined into a larger group (B + C) of mild and moderate dementia. Although these terms were used primarily to indicate gradation within the total group, later experience has shown that most of the patients would have been similarly grouped by psychiatrists using these terms in their usual clinical sense. There were no cases of very severe

Table I.—*Age and Sex Distribution.*

	Group A.	Group B.	Group C.	Group D.	Group (B+C).
Number aged 50–59	10	10	10	10	20
Males	5	5	5	5	10
Females	5	5	5	5	10
Number aged 60–69	10	11	12	10	23
Males	5	5	6	5	11
Females	5	6	6	5	12
Whole Group, aged 50–69 :					
Average age in years	59·4±4·51	60·4±5·33	60·2±5·16	59·3±5·54	60·3±5·10
Number of males	10	10	11	10	21
„ females	10	11	11	10	22

* Abstracted, with permission of the University of Sheffield, from a thesis submitted for the degree of M.D.

or extreme dementia, as no patient was included who clinically displayed any deterioration of vocabulary functions. Great care was taken to select only patients who were fully co-operative and whose failures could be solely ascribed to dementia. Patients with invalidating disorders of the mood or stream of mental activity, or with invalidating physical disabilities, such as poor eyesight, poor hearing or tremor, were rigorously excluded. The age group 50 to 69 was chosen, as this is the period when dementia is commonest, and the patients in each group were equally balanced between the ages 50 to 59 and 60 to 69. The sexes also were equally represented in each group. Although the number in each group was small, this disadvantage was thought to be compensated by the careful selection of patients as to age and sex and co-operation, and so that mental tests could be legitimately assumed to fulfil their intended purpose—matters which many previous investigators appear to have insufficiently regarded.

Tests.—The tests used were :

(1) 1916 Terman Vocabulary test, in full, scored as Babcock (1930).
(2) Form L of the 1937 Stanford Binet test. This was re-scored by discounting credits for vocabulary and appropriately adjusting the credits for the remaining subtests in each year, thus providing a vocabulary-free score.
(3) Babcock test : The original 30 item version (Babcock, 1930) was given, and the 20 items constituting the short form (Babcock, 1933) marked separately, so enabling the calculation of two indices.
(4) Alexander's Passalong test.
(5) Kohs' Block test, original 17 item form.
(6) Porteus mazes, scored up to a maximum of M.A. 16 years (Vernon, 1937).

The dementia ratings were made clinically before testing with the help of the medical officer and nurses in charge of the patient. They were based on informal interviews and, particularly, on observation of general conduct in the ward.

All the tests were personally administered, and split into sessions so that no single session exceeded about an hour.

Results.

Scatter and irregularity.—Two measures of scatter were used—age-level and Pressey scatter. These were calculated for both Form L and the adjusted vocabulary-free Stanford Binet scales (Tables II, III and IV). For both measures there were statistically significant differences between scatter on Form L and on the vocabulary—free scale in groups (B + C) and D, but not in group A. The clear inference is that success on the vocabulary items significantly increased scatter in groups (B + C) and D, but not in group A. From this follows that ability in non-vocabulary items more closely approached vocabulary level in group A than in the other groups, or to express the same facts in another way, there was less discrepancy between the ability on vocabulary and non-vocabulary items in group A than in the other groups—a confirmation of the observations which underlie the Babcock and similar tests.

The measures of scatter proved disappointing. Only one of the group differences was statistically significant—that between groups (B + C) and D for Pressey scatter on the vocabulary free scale. There were indefinite signs that scatter tended to be greater in the intermediate group than at either pole. A contributory reason may be that the tendency for the Form L vocabulary items to exaggerate the true vocabulary ability (Atwell, 1939 ; Brody, 1940) was greater at levels commonly found in group (B + C) than at those in group D, while in group A true vocabulary level was too high for this factor to be important.

Both measures of scatter were higher than expected—probably because of differences between the 1937 scale and the older 1916 Terman revision. Clearly, 8 or more years of age-level scatter and 40 or more points of Pressey scatter—the figures suggested by Harris and Shakow (1937)—cannot be accepted as unequivocal evidence of serious dementia when the 1937 scale is used. Nor do the present results justify the selection of new critical levels. The significance of scatter remains uncertain.

Total irregularity, measured by the range of percentage discrepancies* over all

* See note to Table III.

https://doi.org/10.1192/bjp.88.371.317 Published online by Cambridge University Press

TABLE II.—*Comparison of Mean Age-Level Scatter and Mean Pressey Scatter on Form L and Vocabulary Free Stanford Binet, with D/S.D.diff. where this Exceeds 2·00.*

	Age-level scatter.				Pressey scatter.			
	Form L.	Vocab. free St. B.	L— vocab. free St. B.	Diff./ S.D. diff.	Form L.	Vocab. free St. B.	L— vocab. free St. B.	Diff./ S.D. diff.
Group A	9·95±1·32	9·15±2·64	0·80	...	76·85±31·83	70·84±29·21	6·01	...
Group (B+C)	10·28±1·40	9·21±1·94	1·07	2·94	100·66±37·51	77·26±36·16	23·40	2·96
Group D	10·40±1·43	9·15±1·60	1·25	2·61	95·15±29·39	59·31±24·30	35·84	4·34
Group B	10·43±1·25	9·57±1·72	0·86	...	107·00±31·83	81·12±35·71	25·88	2·47
Group C	10·14±1·60	8·86±2·14	1·28	2·23	94·64±46·26	73·57±37·86	21·07	...

TABLE III.—*Mean Scatter and Irregularity.*

	Group A.	Group (B+C).	Group D.	Group B.	Group C.
Age-level Scatter:					
Form L	9·95±1·32	10·28±1·40	10·40±1·43	10·43±1·25	10·14±1·60
Vocab. free St. B.	9·15±2·64	9·21±1·94	9·15±1·60	9·57±1·72	8·86±2·14
Pressey Scatter:					
Form L	76·85±31·83	100·66±37·51	95·15±29·39	107·0±31·83	94·64±46·26
Vocab. free St. B.	70·84±29·21	77·26±36·16	59·31±24·30	81·12±35·71	73·57±37·86
*Irregularity**	95·0±24·14	106·7±15·52	113·7±15·58	101·86±12·75	111·36±16·89

* Irregularity = Range of "percentage discrepancies" over all items excluding "naming objects." "Percentage discrepancy" = the discrepancy between score on each item and vocabulary norm, expressed as a percentage of the latter.

TABLE IV.—*Differences* in Mean Age-Level Scatter, Mean Pressey Scatter, and Mean Irregularity, with D./S.D.diff. where this exceeds 2·00.*

	A − (B + C).	(B + C) − D.	A − D.	A − B.	B − C.	C − D.	A − C.	B − D.
Age-level Scatter:								
Form L	−0·33 :	−0·12 :	−0·45 :	−0·48 :	0·29 :	−0·26 :	−0·19 :	0·03 :
Vocab. free St. B.	−0·06 :	0·06 :	0 :	−0·42 :	0·71 :	−0·29 :	0·29 :	0·42
Pressey Scatter:								
Form L	−23·8 : 2·61	5·51 :	18·30 :	−30·15 : 3·02	12·36 :	−0·51	−17·79 :	11·85 :
Vocab. free St. B.	−6·42 :	17·95 : 2·32	11·54 :	−10·28 :	7·55 :	14·27 :	−2·73 :	21·81 : 2·30
Irregularity	11·7 :	−7·00 :	−18·7 : 2·91	−6·86 :	9·50 :	−2·34 :	−16·36±2·52	−11·84 : 2·66

* A minus sign (−) in front of the result means that the second of the two groups has the higher value. Some negative values were inevitable whatever the order of presenting the group differences.

tests and subtests, excluding " naming objects," showed more regular tendencies. Group A was significantly less irregular than group D. Scores of above 130 are almost certainly indicative of severe dementia. Such scores, though quite possible on scales including items as easy as " naming the days " and as difficult as " retention of paragraph," are not encountered often enough for the measure to be deemed valuable.

Babcock test and vocabulary minus vocabulary-free Stanford Binet score.—The most commonly used measures of dementia are the discrepancy between scores on vocabulary and on a vocabulary-free Stanford Binet scale, and the Babcock index.* Both appear to be valid in that the results were found to parallel clinical estimations (Tables V and VI). If absolute discrepancies are used, the most reliable measure appears to be the 1930 index. Vocabulary – Stanford Binet discrepancy† discriminated better than the 1933 index between groups A and D, but less well between the intermediate group and group D, while their powers of discriminating between groups A and (B + C) were about equal. At the age of the patients in the present series—the age when dementia most often occurs—there is not a little overlapping of the groups. A critical level below which dementia is probable, if fixed low enough to exclude most non-demented patients, inevitably also excludes some demented patients. On the other hand, most mental hospital dements are worse than those in group D. I therefore believe that it is safer to fix the critical levels rather lower than is generally advised. For vocabulary – Stanford Binet discrepancy I suggest a critical level of 5 M.A. years. Discrepancy of less than 3 M.A. years constitutes good evidence against dementia, while discrepancy of between 3 and 5 M.A. years generally indicates mild or moderate dementia, but with less certainty than scores on either side of these levels. The corresponding indices I suggest for the 1930 Babcock scale are (1) better than − 2·50 ; (2) − 2·50 to − 4·50 ; (3) worse than − 4·50 ; and for the 1933 version, (1) better than − 3·00 ; (2) − 3·00 to − 5·00 ; (3) worse than − 5·00.

These measures are made more sensitive by expressing the discrepancies as percentages of the vocabulary scores (Tables V and VI). The best test now appears to be the 1933 index, followed by the 1930 index, vocabulary – Stanford Binet discrepancy being the poorest. I suggest the following critical levels : for vocabulary – Stanford Binet discrepancy (1) better than 20 per cent. ; (2) 20 to 30 per cent. ; (3) worse than 30 per cent. ; for the 1930 index, (1) better than 15 per cent. ; (2) 15 per cent. to 27 per cent. ; (3) worse than 27 per cent. ; for the 1933 index, (1) better than 20 per cent. ; (2) 20 per cent. to 35 per cent. ; (3) worse than 35 per cent.

Further, critical levels below which dementia is probable may also be established for the various subtests of the Babcock scale. Much reliance, however, should not be placed on the result of a single brief test. The groups overlapped to different degrees in different tests. Thus, some items are more valuable than others (Table VII).

Non-verbal tests.—In a similar fashion, critical discrepancies may be established for the Kohs, Porteus and Passalong tests. Here, too, the result of a single test should not be too much relied on. In such tests particularly the qualitative aspects of the performance are often more revealing than the quantitative result. Moreover, the group overlap was quite large. Indeed, none of the group differences on the Passalong test was significant. In addition, the low " ceiling " of the Porteus test diminishes the value of the vocabulary—Porteus discrepancy. Further, many non-verbal tests are poorly discriminative above M.A. 12. For these reasons I believe that discrepancy scores based on such tests are not very valuable. I suggest a critical level below which dementia is probable of 7 M.A. years or about

* The Shipley test (Shipley, 1940*a* and *b* ; Shipley and Burlingame, 1941) appeared too late to be given to these patients. It has the great merits of simplicity and brevity, and will probably be deservedly popular. As an individual test, however, it is doubtful both whether the vocabulary half affords the subject as good opportunities as a verbally given vocabulary test to score maximally, and whether the other half gives as good opportunities as do more varied tests for the examiner to estimate qualitative aspects of the patient's performance ; and as a group test, it is open to the further criticism, applicable to all such tests in psychotic subjects, that the cause of failure, which may not be intellectual, is not observed.

† This is calculated by subtracting the M.A. score on the vocabulary free Stanford Binet scale from the vocabulary M.A. score.

TABLE V.—*Mean Scores on Various Methods for Measuring Dementia, with S.D.*

	Group A.	Group (B + C).	Group D.	Group B.	Group C.
A	28·75±22·55	48·40±24·83	57·25±22·50	50·80±23·15	46·00±27·34
B	15·85±12·68	24·98±11·77	33·70±10·65	25·67±9·96	24·30±13·50
C	−1·94±1·59	−3·32±1·60	−4·42±1·79	−3·35±1·61	−3·28±1·66
D	−2·64±1·68	−4·06±1·56	−5·26±2·02	−4·04±1·44	−4·08±1·83
E	12·00±11·02	22·14±10·50	31·45±12·31	22·24±10·34	22·05±11·13
F	16·30±10·48	26·56±10·67	37·30±13·87	26·14±9·82	26·95±11·87
G	17·60±13·68	29·26±12·35	48·55±14·27	26·91±13·72	31·50±11·05
H	18·60±13·59	31·05±13·10	48·60±15·03	28·67±14·36	33·32±12·19
I	20·45±15·04	34·86±12·20	52·65±15·54	32·43±14·45	37·17±12·42
J	21·80±15·84	38·33±14·01	54·25±18·83	35·00±13·77	41·27±11·49
K	19·50±12·57	33·16±12·94	47·95±15·62	30·19±12·69	36·00±13·13
L	16·10±11·67	27·19±11·26	46·25±13·24	25·81±11·53	28·50±11·75

TABLE VI.—*Differences in Mean Scores on Various Methods for Measuring Dementia with D/S.D.diff. where this Exceeds 2·00.*

	A − (B + C).	(B + C) − D.	A−D.	A − B.	B − C.	C − D.	A − C.	B − D.
A	19·65 : 3·12	9·85 :	28·50 : 4·00	22·05 : 3·09	−4·8 :	11·25 :	17·25 : 2·25	6·45 :
B	9·13 : 2·72	8·72 : 2·93	17·85 : 4·82	9·82 : 2·75	−1·37 :	9·40 : 2·51	8·45 : 2·09	8·03 : 2·48
C	1·38 : 3·20	1·10 : 2·35	2·48 : 4·64	1·41 : 2·83	−0·07 :	1·14 : 2·14	1·34 : 2·66	1·07 :
D	1·42 : 3·19	1·20 : 2·35	2·62 : 3·36	1·40 : 2·39	0·04 :	1·18 :	1·44 : 2·65	1·22 : 2·18
E	10·14 : 3·45	9·31 : 2·92	19·45 : 5·27	10·24 : 3·03	−0·19 :	9·40 : 2·60	10·05 : 2·93	9·21 : 2·60
F	10·26 : 3·63	10·74 : 3·50	21·00 : 5·87	9·84 : 3·18	0·81 :	10·35 : 2·59	10·65 : 3·44	11·16 : 3·13
G	11·66 : 3·26	19·29 : 5·21	30·95 : 7·00	9·31 : 2·17	4·59 :	17·05 : 4·30	13·90 : 3·60	21·64 : 4·94
H	12·45 : 3·42	17·55 : 4·49	30·00 : 6·63	10·07 : 2·31	4·65 :	15·28 : 3·60	14·72 : 3·68	19·93 : 4·35
I	14·41 : 3·67	17·79 : 4·42	32·20 : 6·66	11·98 : 2·60	4·74 :	15·48 : 3·54	16·72 : 3·91	20·22 : 4·31
J	16·53 : 4·00	15·92 : 3·37	32·45 : 5·89	13·20 : 2·84	6·27 :	12·98 : 2·66	19·47 : 4·53	19·25 : 3·72
K	13·66 : 3·86	14·79 : 3·69	28·45 : 6·24	10·69 : 2·65	5·81 :	11·95 : 2·67	16·50 : 4·07	17·76 : 3·99
L	11·09 : 3·55	19·06 : 5·58	30·15 : 7·69	9·71 : 2·68	2·71 :	17·75 : 4·58	12·40 : 3·43	20·44 : 5·27

KEY TO TABLES V AND VI.

A = Vocabulary *minus* vocabulary-free Stanford Binet (months). B = Above value as percentage of vocabulary score. C = 1930 Babcock Index. D = 1933 Babcock Index. D = 1930 Babcock Index as percentage of norm for vocabulary. F = corresponds to "E" but for 1933 Babcock Index. G = Percentage discrepancy* between Babcock vocabulary norm and average of Babcock subtests 3, 9, 21 and 24. H = as "G," but for subtests 3, 8, 9, 21 and 24. I = as "G," but for subtests 3, 9, 17, 21 and 24. J = as "G," but for subtests 3, 8, 9, 16, 17, 21 and 24. K = as "G," but for subtests 3, 8, 9, 14, 16, 17, 19B, 19C, 21 and 24. L = as "G," but for subtests 3, 9, 21, 22 and 24—*Author's discrepancy score.*

* See footnote to Table III.

50 per cent. of the vocabulary score. Many demented persons will not fall as low as this.

"Discrepancy Score."

Although discrepancy scores based on either the Standard Binet or Babcock scale are valid and useful, to obtain them is, in both cases, an unfortunately long process. The vocabulary test alone cannot be applied in less than 15 minutes, usually takes half an hour, and I have known it take as long as 45 minutes. The remainder of the Stanford Binet scale (1937 version) takes between one and two hours, and the subtests of the Babcock scale—even in the shorter 1933 form—take between 50 minutes and one and a half hours. Moreover, the more demented the patient, the longer the time required; and the longer the time required, the greater likelihood of the entry of invalidating factors, such as loss of attention and co-operation. I confess, too, that personally, I find both scales rather tedious to administer—a complaint I have heard expressed by others—and there is danger of administrative slackness inducing slackness in the patient.

I therefore attempted to devise a method for measuring dementia that would avoid these difficulties. My first task was to devise a method for applying the vocabulary test that would save time, but ensure results close to that of the original technique. The method I evolved (Brody, 1941) requires only 3 to 7 or 8 minutes to apply, and it can be said that the more demented the patient, the greater economy in time. I now re-scored the vocabulary records of the patients in the present series according to the new technique and found the percentage discrepancies for various combinations of subtests of the Babcock scale (Tables V and VI). It can be seen that the significance of the differences between the three main groups was greater than with either of the Babcock indices or the vocabulary – Stanford Binet discrepancy; and also that group B surpassed group C, although the differences are not statistically significant. The most useful combination was found to consist of subtests 3, 9, 21, 22 and 24. I believe that this "discrepancy score," as I have ventured to name it, will be found to be a quickly obtained and sensitive measure of dementia.

The technique for measuring the "discrepancy score" will now be described.

Part A.—The 1916 Stanford Binet Vocabulary Test, applied with the short technique (Brody, 1941).

Part B.—In introducing Part B words like "test," "memory," and "intelligence" should be avoided. The susceptibility of most demented patients to flattery should be used to achieve maximum co-operation. A humble request for the patient "to do something for me if you will, please," or "to have a go at these" usually evokes a good response. I prefer not to mention speed in the general introduction, but merely emphasize the need for it when presenting tests 1, 2 and 3. Otherwise patients sometimes sacrifice accuracy to speed in tests 4 and 5. Praise, encouragement, deprecation of failure and informality are helpful, and may be more freely used than with most tests, as the object is to have the patient do his best. The total time required is usually 15 to 20 minutes.

1. *Opposites* (1930 Babcock Subtest 24).—Emphasize the need for speed. "I want you to do something to show how quick you can be. I'm going to say a word, and I want you to say the *opposite* of the word as quickly as you can think of it. . . . Let's try it. . . . What is the opposite of 'black'? . . . That's it. . . . What is the opposite of 'rich'?" Correct if necessary. If the subject fails to understand "opposite," he can sometimes comprehend the task by explaining "You can have 'black' and 'white,' and you can have 'rich' and 'poor,' etc." . . . "Now don't forget that you have to say the *opposite* of the word, and as quickly as you can think of it."

When the testee has understood the task, test in order the opposite of—

1. up.
2. bad.
3. in.
4. slow.
5. day.
6. thick.
7. innocent.
8. generous.
9. cruel.
10. dangerous.
11. rare.
12. ancient.
13. prohibit.

Score.—Time each response.

Credit 2 for each correct response given in 1 sec. or less.
" 1 " " " " from 1 to 2 secs.
" ½ " " " " " 2 to 10 secs.
" 0 if the response is wrong or not given within 10 secs.

https://doi.org/10.1192/bjp.88.371.317 Published online by Cambridge University Press

TABLE VII.—*Lower Limit of Discrepancies Indicative of Serious Dementia on Various Subtests of the Babcock Scale.*

Subtests.	Points discrepancy.	Percentage discrepancy.
2 ; 4 ; 18 ; 19A ; 24	— 1·5	10
1 ; 5 ; 6	— 3·0	20
19B ; 19C ; 20 ; 22 ; 23	— 4·5	30
7 ; 12 ; 14 ; $15C_3$; 21	— 6·0	40
8 ; 10 ; 15B ; $15C_1$; $15C_2$	— 7·25	50
3 ; 9 ; 11	— 9·0	60
$15C_4$; 16 ; 17	— 9·5	65

TABLE VIII.

Word score .	78.	74–77.	70–73.	65–69.	60–64.	54–59.	49–53.	44–48.	38–43.
M.A. . .	20.	19.	18.	17.	16.	15.	14.	13.	12.
1·0	94	94	94	94	93	93	93	93	91
2·0	88	87	87	87	87	86	85	85	84
3·0	83	81	81	81	80	80	78	78	76
3·5	80	78	78	78	77	77	75	74	72
4·0	77	75	75	75	74	73	71	70	68
4·25	75	74	73	73	73	71	69	68	66
4·50	74	72	72	72	71	70	67	67	64
4·75	72	71	70	70	70	68	66	65	61
5·00	71	69	69	69	68	67	64	63	60
5·25	70	68	67	67	66	65	62	61	58
5·50	68	67	66	66	65	63	61	59	56
5·75	67	65	64	64	63	61	59	57	54
6·00	65	63	62	62	61	60	57	55	51
6·25	64	62	61	61	60	58	55	54	50
6·50	63	60	59	59	58	56	53	52	48
6·75	61	58	58	58	57	55	51	50	46
7·00	60	57	56	56	55	53	50	48	44
7·25	58	55	55	55	53	51	48	46	42
7·50	57	54	53	53	52	50	46	44	40
7·75	55	52	51	51	50	48	45	42	38
8·00	54	51	50	50	49	47	43	40	36
8·25	53	49	48	48	47	45	41	39	34
8·50	51	48	47	47	46	43	39	37	32
8·75	50	47	45	45	44	41	37	35	30
9·00	48	45	44	44	42	40	35	33	28
9·25	47	43	42	42	41	38	34	31	26
9·50	45	42	40	40	39	37	32	29	24
9·75	44	40	39	39	37	35	30	27	21
10·00	42	39	37	37	36	33	29	26	20
10·25	41	38	36	36	34	32	27	24	18
10·50	39	36	34	34	33	30	25	22	16
10·75	38	34	33	33	31	28	23	20	13
11·00	37	33	31	31	29	27	21	18	12
11·25	35	31	30	30	27	25	20	16	9
11·50	34	30	28	28	26	23	18	14	8
11·75	32	28	27	27	24	22	16	13	6
12·00	31	27	25	25	23	20	15	11	4
12·25	29	25	23	23	21	18	12	9	2
12·50	28	23	22	22	20	17	11	7	0
12·75	27	22	20	20	18	15	9	5	+3
13·00	25	20	19	19	16	13	7	3	+5
13·25	24	19	17	17	15	12	5	2	+7
13·50	22	17	16	16	13	10	4	0	+9
13·75	21	16	14	14	11	8	2	+3	+11
14·00	20	14	12	12	10	7	0	+5	+13
14·25	18	13	11	11	8	5	+2	+6	+15
14·50	17	11	9	9	6	3	+4	+8	+17
14·75	15	10	8	8	5	2	+5	+10	+19
15·00	14	8	6	6	3	0	+7	+12	+21
15·25	12	7	5	5	1	+2	+9	+14	+23
15·50	11	5	3	3	0	+3	+11	+15	+25
15·75	10	4	2	2	+1	+5	+12	+18	+27
16·00	8	2	0	0	+3	+7	+15	+20	+29

If the score on the list of thirteen words is zero, credit 1 for each sample word correctly answered.

II. *General information* (1930 Babcock Subtest 3).—Again emphasize the need for speed. "Now here's another to show how quick you can be. I'm going to ask you some questions, and I want to see how quickly you can answer them. You don't have to be clever to answer them—it's just that maybe they're about things you have come across, or maybe not. You'll probably know some of the answers although you may not know them all. What I want you to show me is how quickly you can think of things that you really know. When I've asked the question, think hard, and tell me the answer as quickly as you can. Now—"

a. What is your first name ?
b. Are you married ?
c. How many brothers (or sisters) have you ?
} Given to test comprehension of the task. Time, but do not score.

1. How old are you ?
2. Which king is now on the throne ?
3. What is the name of the Prime Minister ?
4. What is the name of the President of the United States ?
5. Who is the Mayor of . . . . . ? (If in a rural area, or there has been a change of mayor since the patient was admitted to hospital, substitute "What is the name of the leader of the Labour (Opposition) party ?)
6. What sort of things are made in Birmingham ? (Any metallic object, correct. Any non-metallic object, e.g. clothing, wrong.)
7. About what was the last amendment to the constitution of the United States ? (This is the question in the original Babcock test and, is reproduced here to represent the sort of question required. The content may be about any recent event of national importance, e.g. (Nov., 1941) the Lease-Lend Act.)

Score.—Time each response.

Credit 3 for each correct response given in 1 sec. or less.
,, 2 ,, ,, ,, ,, from 1 to 2 secs.
,, 1 ,, ,, ,, ,, 2 to 10 secs.
,, 0 if the response is wrong or not given within 10 secs.

III. *Substitution test* (1930 Babcock Subtest 9).—Again emphasize the need for speed. Forms as in the Babcock test are used. "You see these drawings" (indicate the whole sheet). "They are all like these up here" (indicate key). "But these (indicating key) have numbers in them. What I want you to do is to write in each of these (indicating blanks) the same number that the same shape has up here (indicating key). For example, this would be number . . ." (Test with the first drawing in the first line, and with two others at random but not in the first line as this makes the first line too easy.) . . . "That's right. Start here (indicating first of the first line) and fill them all in in turn as you come to them. Do them as quickly as you can, but be sure to get them right—as *quickly* as you can, but be sure to get them right."

Continue to explain until the task has been understood. If the subject is very slow, estimate, after five minutes, the time that he would take to complete the task. Small differences in time are not important at low functional levels.

Score :

Time (seconds).	Points.	Time (seconds).	Points.
48–49	21	107–110	11½
50–51	20½	111–114	11
52–53	20	115–123	10½
54–55	19½	124–131	10
56–57	19	132–140	9½
58–59	18½	141–149	9
60–62	18	150–159	8½
63–66	17½	160–168	8
67–69	17	169–205	7½
70–74	16½	206–224	7
75–80	16	225–230	6½
81–85	15½	231–241	6
86–90	15	242–518	5½
91–93	14½	519–600	5
94–95	14	4 correct in 5 min.	4
96–97	13½	3 ,, ,,	3
98–99	13	2 ,, ,,	2
100–103	12½	If get idea	1
104–106	12	No comprehension	0

Add $\frac{1}{50}$ to the time for each error.

IV. *Sentence repetition* (1930 Babcock Subtest 22).—"Now in this one, you don't need to be especially quick. All I want you to do is to say exactly what I say. I'm going to say some

https://doi.org/10.1192/bjp.88.371.317 Published online by Cambridge University Press

sentences, and when I finish each one, I want you to say exactly the same. They are not tongue-twisters. This is just to see how long a sentence you can say. Listen carefully and get them *exactly* right." Start most subjects with sentence 4. In the absence of failure, alternate sentences may be given up to number 8.

1. Mamma.
2. Nice kitty.
3. In summer the sun is hot.
4. When the train passes, you will hear the whistle blow.
5. Last year, fifteen new houses were built on this street.
6. The price of potatoes is fifteen pence per peck.
7. The road up the mountain was very rough and steep.
8. Walter had a fine time on his holidays; he went fishing every day.
9. The river flowed by with a pleasant soothing sound.
10. The breeze was strong, and the sailboat sped over the water.
11. The cricket match yesterday was won in the very last minute.
12. To take a picture of the moon, one must have a large camera and a good lens.
13. At the level crossing, the engine driver will blow the whistle and the fireman will ring the bell.
14. Children, it is necessary to work for a living; you must go to your school every morning.
15. An old man sitting in the corner put down his paper and looked at him as he entered the room.
16. Last year, there were only twenty children in the school, but this year, there are fifty as the town has grown so.
17. When the girl entered the tram-car with her bag, all the seats were filled and people were standing in aisles.
18. The church was cool and dark, and at the far end there was a high altar of marble.
19. Walter likes very much to go on visits to his grandmother because she always tells him many funny stories.
20. Yesterday, I saw a pretty little dog in the street; it had curly brown hair, short legs and a long tail.
21. As the great red sun came over the hills, the Indians broke camp and prepared for another day's hard work.
22. The clouds hung low in the valley, and the wind howled among the trees as the men went on through the rain.
23. One thing a nation must have to become rich and great is a large secure supply of wood.

Score:

Credit 1 for each sentence correctly repeated.

Credit ½ for each sentence, up to No. 15, repeated with only one error. After No. 15, no half credits are allowed and the repetition must be entirely correct.

Minor errors may be overlooked in the earlier sentences (e.g. " pence *a* peck " for " pence *per* peck " in No. 6), but not in the later ones (e.g. " full " for " filled " in No. 17).

V. *Designs from memory* (1930 Babcock Subtest 21).—Use the designs for test 6 of the Performance Scale of the Army Mental Tests, as in the Babcock test.

For designs " a " and " b " (in turn): " On this sheet (demonstrate *blank* side) there's a little drawing. I'm going to show it to you for ten seconds. Then I'm going to take it away, and I want you to draw it from memory. Are you ready? Look at it very carefully so that you will remember it."

For designs " c " and " d ": " On this sheet there are *two* little drawings. I'm going to show you them for ten seconds like the others, then I'm going to take them away, and I shall want you to draw *both* drawings. So, don't forget to look carefully at *both* drawings . . . *both* drawings."

After 4 seconds' exposure, repeat " *Both* drawings."

If, on hearing the instructions, the subject objects that " he was never any good at drawing," he should be assured that drawing skill is unnecessary, but without hinting at the value of observing the structure of the designs. If, as sometimes happens, a patient wishes to start drawing after only two or three seconds' exposure, he should be urged to " keep looking." The patient should not be given the pencil until after the full exposure of the first drawing, lest, having misunderstood the instructions, he starts to copy the design before it has been removed. Experience with the first design rarely fails to prevent a recurrence of this error.

Score.—Credits are awarded as in the Army Performance Scale. If the subject fails to obtain an " Army Credit," credit as follows:

1 point each if two lines are crossed in design " a," or if there are four squares.

½ for each outer shape in " c.1 " and " c.2."

1 for each outer shape in " b," " d.1 " and " d.2."

Transform these credits into points in accordance with this table:

Credits . .	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Points . .	5½	6½	8	8½	9	10	10½	11	12	13	14	15	16½	17½	19	20	21

Calculation.—To find the " discrepancy score," calculate the average point score for the five subtests of part B, subtract it from the norm for vocabulary, and express this difference as a

percentage of the vocabulary score. Most scores commonly encountered can be read off from Table VIII at the point where the individual's average point score row intersects his vocabulary column. Norms are the same as for the 1933 Babcock test, thus—

Vocabulary M.A.:

20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5
Norm.:															
17·4	16·4	16	15·8	15·5	15	14	13·4	12·4	11·6	10·6	9	8	7	6	5

Interpretation.—Suggested interpretations are:

Better than 25 = Probably not, or doubtfully demented.
Worse than 40 = Probably seriously demented.
25 to 40 = Mildly or moderately demented.

It is hoped that further results will permit finer and more reliable gradations as well as data for the application and interpretation of this test in other age groups. Obviously these scores and norms are for the average of all 5 subtests, and if any are omitted, the results will not be exactly comparable with those of the full scale.

Discussion.—Although as a research instrument, this test has all the faults of the " hotch-potch " so scathingly criticized by Cattell (1940), it has a number of advantages for practical clinical work. It is short and simple to administer and score, but varied enough to prevent boredom and to allow observation of the quality of the patient's performance. Requiring no elaborate apparatus, it can be informally introduced into routine clinical interviews, if necessary, in two sessions. No subtest contains material embarrassing to adults. Since devising it I have used it in practice with promising results.

I must conclude with a warning. Though intended to assist in the diagnosis and measurement of dementia, the scale will not provide automatic diagnosis. The examiner must decide in each and every case the validity of using it and the meaning of the result. To quote Earl (1940), " Mental tests are clinical instruments whose value, like that of all clinical instruments, depends to a great extent on the ability of the user. Always the quantitative results obtained from them must be interpreted in the light of the qualities of behaviour observed during the test, and eventually in the light of all the available data."

Summary.

1. The difference between scores on vocabulary and on other tests can be used as an adequate measure of dementia. Expression of the discrepancy as a percentage of the vocabulary score increases the sensitivity of this measure. Critical levels above and below which dementia is unlikely and highly probable may be established.

2. Unfortunately in later life, when dementia most commonly occurs, normal variation is so great that critical levels low enough to exclude most non-demented patients also exclude many demented patients.

3. Reliable discrepancy measures are obtained with the Stanford Binet and Babcock scales, for which interpretations of scores are suggested.

4. Both these techniques are, however, long and tedious. A technique is described for calculating the " discrepancy score "—the percentage discrepancy between score on vocabulary and on 5 selected Babcock items, viz. " opposites," " general information," " substitution," " sentence memory " and " designs from memory." The test has the advantages of simplicity and brevity in application and scoring combined with enough variety to prevent boredom and to permit observation of the quality of the patient's performance. Requiring no elaborate apparatus, it can be informally administered, and is acceptable to adults. Experience since the test was devised encourages the hope that it will prove to be a useful clinical instrument in routine psychiatric practice.

I am indebted to Dr. R. Ström-Olsen, Physician Superintendent, for permission to use the clinical material of Runwell Hospital.

References

Atwell, C. R. (1939), *J. educ. Psychol.*, **30**, 467.
Babcock, H. (1930), *Arch. Psychol.*, No. 117.

Idem (1933), *A Short Form of the Babcock Examination for the Measurement of Mental Deterioration*. Chicago: Stoelting.
BRODY, M. B. (1940), *J. ment. Sci.*, **86**, 532.
Idem (1941), *Ibid.*, **87**, 88.
CATTELL, R. B. (1940), *J. educ. Psychol.*, **31**, 161.
EARL, C. J. C. (1940), *J. abnorm. (soc.) Psychol.*, **35**, 428.
HARRIS, A. J., and SHAKOW, D. (1937), *Psychol. Bull.*, **34**, 134.
SHIPLEY, W. C. (1940*a*), *Ibid.*, **37**, 438.
Idem (1940*b*), *J. Psychol.*, **9**, 371.
Idem and BURLINGAME, C. C. (1941), *Amer. J. Psychiat.*, **97**, 1313.
VERNON, P. E. (1937), *Brit. J. educ. Psychol.*, **7**, 72–88 and 115–137.