

PSYCHOLOGICAL TESTS OF RESEARCH SCIENTISTS¹

ANNE ROE

NEW YORK, N. Y.

THIS paper reports some special aspects of a research project which has been carried on for the past three and a half years. The major part of the project was the clinical study of eminent research scientists in various fields of science, and a consideration of differences and similarities among them in life histories and test performances. The full reports of the study have been published elsewhere [2, 3]. This paper presents certain interrelations between the tests used that are of particular interest. An earlier paper presented similar but less complete data on the biologists alone [4].

The tests employed were the Rorschach, the TAT, and a Verbal-Spatial-Mathematical (VSM) test constructed for the purpose by the Educational Testing Service, the three subtests being separately scored. The verbal test is an antonym test; the spatial test requires identification of drawings of the same object seen in different positions, and the mathematical test is chiefly one of mathematical reasoning. The mathematical test was not difficult enough for the physicists, so that they are not included in any correlations using it.

The subjects in the total study included 64 ranking research scientists: 20 biologists, 22 physical scientists, and 22 social scientists (psychologists and anthropologists). For three of these subjects there are no test data available, and there are some deficiencies in the data for the others, as is evident in the tables. These subjects were selected for their eminence in research, in the estimation of their peers. All are men, all but one are American-born, and all received the bulk of their training here and

were still scientifically productive when studied. They range from 31 to 60 in age. Details on selection are given in the monographs, which also contain tables of raw data not reprinted here.

EFFECT OF AGE ON TEST PERFORMANCE

Correlations with age are given for various aspects of test performance in Table 1.² The subtests of the VSM are entered first in the table. Clearly, only the spatial test shows any effect of age, within this age range, but this is great enough to be of some importance.

TABLE 1
EFFECT OF AGE ON TEST PERFORMANCE:
CORRELATIONS

| Test Items | N | Correlation with Age |
|---------------------------------|----|----------------------|
| Verbal Test, number right | 59 | -.108 |
| Spatial Test, number right | 57 | -.397 |
| Mathematical Test, number right | 39 | +.001 |
| Rorschach: R | 61 | -.194 |
| T/R | 59 | -.150 |
| W | 61 | -.068 |
| M | 61 | -.204 |
| ΣC | 61 | -.204 |
| F% | 61 | +.113 |
| F+% | 61 | -.119 |
| ITS | 61 | +.222 |
| TAT: length of story | 61 | -.193 |

Not all the Rorschach determinants are appropriately handled by the usual correlational techniques, but a few which seem of particular interest, and which have wide ranges and reasonably normal distributions, have been included. None of these correlations is significant at the 5 per cent level ($r = .25$ for this sample).

The correlation between age and length of

¹Part of a study financed by a grant from the Division of Research Grants and Fellowships of the National Institutes of Health of the United States Public Health Service.

²All calculations were done by Mr. Lassar Gokin.

TAT story is also not significant. The measure of length is a crude one — the average number of lines of typescript for the stories for each man. These protocols were all recorded in shorthand; there are doubtless some omissions of words or phrases, but the record is reasonably accurate.

It would seem, then, that within this age range, and with the exception of the spatial test, variability in age does not have an important influence on test performance.

VSM TEST INTERCORRELATIONS

These correlations, given in Table 2, are based on raw scores, i.e., on the number of items answered correctly in each subtest. Only the correlation between the verbal and spatial subtests is significant. This is not materially changed if age is partialled out. Descriptions

TABLE 2
VERBAL-SPATIAL-MATHEMATICAL TEST
INTERCORRELATIONS: NUMBER RIGHT

| | Verbal | | Spatial | |
|--------------|----------|----------|----------|----------|
| | <i>N</i> | <i>r</i> | <i>N</i> | <i>r</i> |
| Spatial | 57 | +.332 | | |
| Mathematical | 39 | +.144 | 39 | +.212 |

of work on the spatial test would confirm that it can be done in different ways. This test was a rather unsatisfactory one, inasmuch as its significance is not clear. The distribution of scores was platykurtic, and its correlations with other tests (not used in this study) are confusing.

PRODUCTIVENESS OR OUTPUT

It can be hypothesized that there is a general factor of productiveness or ease and amount of output for the various tests, which might be

manifested in such ways as the amount of work attempted in each test. For the VSM subtests these would be the number of items which the subject had tried to do (computed as the last one marked; if some items were omitted before that it was assumed that he had read and at least considered them), for the Rorschach the total number of responses, and on the TAT the average length of the stories. The correlations are given in Table 3. Two of the VSM correlations are low, but significant, but it is not certain that this is not primarily a speed factor inasmuch as this test was timed and some subjects did not finish all the items. But the correlation between verbal "tried" and Rorschach *R* is also significant, although low. Length of TAT story and number of responses on the Rorschach have a positive correlation of .499, which is the highest in the table.

RELATIONS BETWEEN VSM AND
PROJECTIVE TESTS

Relations between scores on the VSM and certain aspects of projective test performance which seemed relevant were studied by the correlations given in Table 4. Here are also given partial correlations which indicate the relative effects of age and total number of responses on the Rorschach (*R*). Total number of responses on the Rorschach clearly affects the total number of any particular kind of Rorschach response, and with the range in *R* as wide as it is in this group, this must always be considered. The mathematical test shows no significant relations with any of the other elements studied, but it is possible that there is some relation between this and percentage of Rorschach responses using Form alone (*F%*). (Correlation between *R* and *F%* in this group is -.028.) The group is small inas-

TABLE 3
INTERCORRELATIONS OF "OUTPUT"

| | S tried | | M tried | | R | | TAT length | |
|-------------------------|----------|----------|----------|----------|----------|----------|------------|----------|
| | <i>N</i> | <i>r</i> | <i>N</i> | <i>r</i> | <i>N</i> | <i>r</i> | <i>N</i> | <i>r</i> |
| Verbal, no. tried | 57 | +.317* | 39 | +.369* | 59 | +.264* | 59 | +.116 |
| Spatial, no. tried | | | 39 | +.225 | 57 | +.233 | 57 | +.095 |
| Mathematical, no. tried | | | | | 39 | +.098 | 39 | -.237 |
| Rorschach, <i>R</i> | | | | | | | 61 | +.499* |

**P* = .05 or less.

TABLE 4
CORRELATIONS BETWEEN NUMBER RIGHT ON VSM SUBTESTS AND VARIOUS PROJECTIVE TEST MEASURES.
PARTIAL CORRELATIONS WITH AGE AND R.

| 1 | 2. Verbal Right ($n = 59$) | | | 2. Spatial Right ($n = 57$) | | | 2. Math. Right ($n = 39$) | | |
|------------|------------------------------|--------------------|------------------|-------------------------------|--------------------|------------------|-----------------------------|--------------------|------------------|
| | r_{12} | $r_{12 \cdot age}$ | $r_{12 \cdot R}$ | r_{12} | $r_{12 \cdot age}$ | $r_{12 \cdot R}$ | r_{12} | $r_{12 \cdot age}$ | $r_{12 \cdot R}$ |
| Rorschach | | | | | | | | | |
| R | +.279* | +.265* | | +.048 | -.030 | | -.074 | -.075 | |
| ITS | -.007 | +.018 | -.110 | +.054 | +.159 | +.040 | -.213 | -.219 | -.204 |
| W | +.153 | +.146 | -.053 | +.062 | +.108 | +.040 | -.045 | -.046 | +.014 |
| M | +.221 | +.204 | +.322* | +.015 | -.073 | -.020 | -.137 | -.140 | -.133 |
| F% | -.003 | +.009 | +.005 | -.000 | +.049 | +.002 | +.234 | +.235 | +.263 |
| F + % | -.137 | -.152 | -.134 | +.340* | +.321* | +.337* | +.199 | +.201 | +.207 |
| ΣC | -.304* | -.337* | -.893* | -.025 | -.128 | -.103 | -.021 | -.022 | +.066 |
| TAT | | | | | | | | | |
| Length | -.545* | -.580* | | -.227 | -.254 | | +.145 | +.148 | |

* $P = .05$ or less.

much as the physicists did not take this test.

The spatial test shows significant relations only with the measure of the adequacy of form on the Rorschach ($F + \%$), and age and R are apparently irrelevant.

The verbal test is positively correlated with total numbers of responses on the Rorschach and with number of human movement responses (M) when the effect of total responses is discounted. The negative correlation with sum of color responses (ΣC) is very high when R is partialled out. This is a rather consistent picture — the verbally skilled person, productive on the Rorschach, and rather introversive. The strong interest in persons shown also by the high M is also to be noted. I have elsewhere speculated that the reliance upon verbal modes of thinking and the high verbal skills demonstrated by individuals who are much concerned with personal relations may be due to the fact that words are our chief means of interpersonal communication, and those to whom this is important develop verbal facility. The negative relation with sum of color responses would indicate a much stronger degree of introversiveness than seems probable in this total picture. This is very likely due to the manner in which ΣC is computed. The FC responses are very common in those with high M [5, and shown also in my data], but they are weighted only half as strong as the CF responses; hence, the picture may be distorted. Interrelations within Rorschach elements themselves are not appropriately considered here, but they will be dis-

cussed at length in another paper.

The negative correlation between verbal right and TAT length is somewhat difficult to interpret. The correlation between verbal items tried and TAT length was negligible, although positive. Perhaps the point is that precision in the use of words is not related to fluency in their use.

SHADING SHOCK AND VSM SCORES

The commonly accepted interpretation of shading shock as a pervasive, basic anxiety leads to speculation as to whether the presence of such anxiety has a disrupting effect upon intellectual performances of various sorts. If so, it might be manifested in lowered scores, or it might be manifested in greater scatter on test scores. A simple test is insufficient in so complicated a situation, but it seemed worth while to examine these data in this respect. The results are shown in Table 5, which gives

TABLE 5
SHADING SHOCK ON RORSCHACH AND VSM SCORES
AND SCATTER

| VSM Scores | Shading Shock | | |
|---------------------------------|--------------------|------------------|--------------------|
| | Strong $N = 16$ | Mild $N = 23$ | Absent $N = 22$ |
| Verbal right, ave. | 53.5 | 58.0 | 60.1 |
| Spatial right, ave. | 11.1 | 9.8 | 11.9 |
| Mathematical right, ave. | 12.9 | 14.0 | 16.8 |
| V-S, Scatter in σ scores | 1.0 | 1.1 | 0.7 |

mean V , S , and M scores and mean scatter for the subgroups of those with strong, mild, and no shading shock. The measure of scatter is

the difference between sigma score on verbal and on spatial tests. (The mathematical test was not taken by the physicists.) Although means for verbal and mathematical tests increase regularly with decreasing scatter, analysis of variance gives F 's of 1.00 and 1.20, which are not significant. On these data, then, it cannot be shown that shading shock has any effect on test score.

MISCELLANEOUS COMPARISONS

A few further comparisons, which seemed of interest, have been made between Rorschach and TAT performance.

Length of TAT stories correlates + .555 with number of whole responses (W) on the Rorschach, which drops to + .339 when R is partialled out, but is still significant. The point of this seems to be that the longer stories are also likely to be more complete, to encompass more, and this partakes of the same attitudes which result in large production of whole responses on the Rorschach.

One would also like to know whether originality on the two tests was related. For various reasons, it seemed preferable to use P (popular responses) on the Rorschach, rather than O (original responses). The measure on the TAT for each subject was the number of unusual stories he gave, plus one-half the number of stories to which he gave a distinctly unusual twist. In one instance, the measure is one of conformity, in the other, of departure from conformity. They should have a negative correlation if they have any basis in common. The correlation is completely insignificant, + .005.

Finally, Piotrowski's remark that the hu-

man movement responses (M) "indicate the testee's preoccupation with the future" stimulated another computation. The extent to which the subjects completed their TAT stories by including a future was extremely varied, even though they were specifically asked for this. To give stories with a future is not necessarily a measure of the extent to which the subject is preoccupied with the future, but it would seem probable that there is some relation. Therefore, a correlation was run between the number of human movement responses on the Rorschach and the number of stories which included the future. This is + .297, and, although low, is significant at the 2 per cent level.

SUMMARY

This paper has reported some comparisons of test results on a verbal-spatial-mathematical test and the Rorschach and the TAT given to 61 eminent research scientists.

Received March 20, 1951.

REFERENCES

1. PIOTROWSKI, Z. A Rorschach compendium—revised and enlarged. *Psychiat. Quart.*, 1950, **24**, 543-596.
2. ROE, ANNE. A psychological study of eminent biologists. *Psychol. Monogr.*, 1951, **65** (14), No. 331.
3. ANNE ROE. A psychological study of eminent physical scientists. *Genet. Psychol. Monogr.*, 1951, **43**, 121-235.
4. ROE, ANNE. Psychological examinations of eminent biologists. *J. consult. Psychol.*, 1949, **13**, 225-246.
5. WITTENBORN, J. R. Statistical tests of certain Rorschach assumptions: The internal consistency of scoring categories. *J. consult. Psychol.*, 1950, **14**, 1-19.