psychologists who carry out research must be endowed with the necessary patience and care so vital in producing accurate and meaningful results.

**Time for research**

One often hears it said by reproachful psychologists, in practice, that there is simply not enough time to carry out research as part of their work. This is usually based on too large a case load, a shortage of psychologists in a particular area and/or the possible opposition or lack of encouragement of certain administrators who employ psychologists, and who consider it an unnecessary luxury to carry out research.

In response to this, one can only say that while some psychologists enjoy using their professional expertise in almost totally practical ways, others seek an added dimension to their work. It is the practical work, combined with applied research, which gives many psychologists an important job satisfaction, as well as providing relevant new information which could be of value to other psychologists in practice.

One should anticipate few objections to a psychologist's carrying out research in areas in which he is already engaged. The object must be to sharpen his diagnostic skills, or to improve therapeutic techniques. Information gained from this type of applied research should be written up and published if possible, so that other psychologists can also benefit from the work done.

It is unfortunate that much research is currently being done by non-practitioners, and much which is carried out has little relevance to practical application. Those psychologists who are interested and capable of carrying out good research should be allowed to do this as an important part of their function.

**Research which has practical value**

Much has already been said about the importance of psychologists' carrying out research in a number of possible ways. Those include:

1. Surveys in order to establish how serious a problem may be, or the prevalence of certain conditions.
2. Devising new, or improving old, diagnostic techniques.
3. Devising and assessing new, or improving old, therapeutic or remedial techniques, in regard to helping children with their learning problems and/or emotional-behavioural difficulties.
4. Researching ways through which teachers and parents can exercise their function more effectively as directors of desirable behaviour in their children.
5. Organizing and assessing the value or effectiveness of therapeutic methods in the neighbourhood and in the community, especially if it is possible to assess the benefits to children and members of families, and the interaction of members of a community.

These are but some of the types of research that can be carried out by psychologists in practice.

At present, all three of the prerequisites which have been discussed, which lead psychologists to do research, vary. Some school psychologists and school psychological services have ample time, or make the time, for research. Others have very little time, if any, for carrying out applied research.

Certainly, the training of psychologists working in schools with pre-school and school children has, and must continue to have, greater breadth and depth. In the past, training in conducting research was limited, whereas today it is an integral part of the training programmes for psychologists. The profession of psychology has thereby been enriched considerably. It has also resulted in the commensurate respect the profession has gained from the general public, and from other members of professions allied to work in mental health. Obviously, the psychologist's professional responsibilities to advise, diagnose, teach and treat should not suffer at the expense of conducting research. On the contrary, the psychologist's interest in and will to carry out research should enhance his usefulness, along with other colleagues, in improving the mental health of persons with whom he is involved.

---

**Raven Coloured Progressive Matrices Intelligence Test in Thailand and in Denmark**

ANDERS POUlsen

Dansk Psykologisk Forlag, Copenhagen, Denmark

It might appear regressive to occupy oneself with, and write about, Raven's 1956 intelligence test. Presumably it is no longer applied at all in Denmark. Danish school psychologists have never considered it a good and reliable test, but it was at one time used for certain screening purposes, and now and then it was included in the 'battery' that had to be applied in attempts to penetrate as far as possible into a child's wholly special problems.

In the course of contacts with colleagues, and with educational institutions in other countries, I have again and again ascertained that Raven's Progressive Matrices (RPM) test is not at all obsolete everywhere. In particular, I have learned that it is applied in certain school systems in some Third World countries as a link in entrance examinations and in separating children into schools' various 'streams'.

As I had available some test results gathered in Thailand almost 20 years ago, which showed that the test was completely inapplicable in the Thai village connection, I saw a certain sense in supplementing this material, and possibly thereby taking part in shaking, just a little, the faith in the rightness of the test for such purposes as those given.

The examination, where all children are tested singly, is here repeated briefly.

**About the test**

J.C. Raven states in the handbook that it is well suitable for 'cross-cultural research', as it is non-verbal and easy to teach. He further stresses that no time limits are set for the test, to ensure that success depends only upon a person's present capacity for intellectual activity.

The test is in many places referred to as 'culture free'. But this statement was shaken as early as 1958 by Burke, and similarly by Anastasi (1970). The test can be said to be, perhaps more than any other test, 'culturally fair' even if not 'culture free', however, and Vernon's book (1969), among others, supports this view. The doubt that continuously exists about the test's qualities is, I suppose, a chief reason why this easily accessible, quickly accomplished and non-linguistic test is being applied more in the Third World than elsewhere and is thus also being used for purposes of drawing serious and far-reaching consequences for children.

**Thailand**

Village. During my work in a north-eastern Thai village in 1961–1962, when the pupils in the school's first grade were the 'sample' on the basis of which we worked, we tested 47 children using RPM. We did not count on the results being useful for telling us anything significant about the children, but considered it reasonable to spend time collecting these
results to obtain knowledge about the tests. According to RPM there were in this first grade neither 'superior' nor 'above average' children, whilst nearly 62 per cent were 'intellectually defective' and 22 per cent 'definitely below average'; results which were completely nonsensical both from our own good knowledge of the children as single individuals and of the teachers' evaluations. The test was therefore completely unsuitable for its given purpose with regard to this grade of pupils in this particular village.

On the other hand, several well-arranged examinations by internationally acknowledged researchers appear to show that, in the Third World countries, the test becomes somewhat more reasonable if it is applied to children who have received some years tuition. On this basis I concluded that it would make sense to test the pupils in the fourth grade — the last obligatory school year — in the same Thai village, and the 51 pupils of the fourth grade level were tested in January 1978.

Here again not a single child was placed in the two groups 'above average' and 'superior', but more than 80 per cent scored points for a placing in the lowest group, 'defective'. The results thus could not confirm that the test was 'better' for children who had received tuition for some years. The results are fully as absurd as the originals from the first grade, evaluated both from the teachers' views regarding intelligence and level in main subjects and from my own good knowledge of the children as individuals.

Bangkok. On the basis of the results given above it might be interesting to establish what the results would be in the large city surroundings of Bangkok.

A school in an average suburb, where the majority of the parents were employed as junior civil servants, clerks, shop assistants and bus and taxi drivers, was selected. Twenty-five children in a first grade and 30 children in a fourth grade were tested.

The results put both grades predominantly in the lower groups, yet there is a quite different picture from that of the village. Although no child scored for 'superior' there were not so few in the group 'above average' as there had been in the village. The results have a more 'normal' distribution at the fourth grade level than at the first grade level, but perhaps it is in these surroundings that the years at school are reflected.

Here, but only in the fourth grade, appears a distinction between boys and girls. The boys were much sharper than the girls at coping with the test. It is, furthermore, the only grade in the whole examination that has shown variance of any significance between boys and girls. For both grades, however, there was no good conformity between test results and teachers' assessments regarding the pupils' capacities and levels in main subjects.

My conclusion for Thailand had to be: the groups are small, and it is not possible, from these results, to give a general assessment of Thailand. However, RPM is not suited to give any authentic picture of the intelligence of the children tested, neither in a village nor in a large city in Thailand. In addition, it would not amaze me to find, should more comprehensive examinations be made, with this test, of children of all ages throughout Thailand, that the test is entirely unsuitable in the Thai culture.

Denmark

It then became appropriate to involve some Danish children in this little examination. Were there to be the merest reasonable possibility of comparing results with the samples from Thailand it would be necessary to find a large city and a country environment with the greatest possible racial difference.

We chose (1) a three-directional school in the northern neighbourhood of Copenhagen, in an area with a fairly varied population basis, and (2) a couple of very small schools in villages in the northwestern part of the country — villages situated fairly remotely from larger cities.

All pupils of first and fourth grades were tested.

The northern region of Copenhagen. In the first grades (60 children) the results were just as extreme as in the Thai village, but with reversed signs. Nearly 85 per cent of the children fell into the groups 'superior' and 'above average', and none into the group 'defective'.

In the fourth grade (56 children) 'only' 41 per cent were in the test's two best groups, and here again none in the lowest.

None of the grade levels had a points distribution even reminiscent of a normal distribution, and for none of the grade levels was there any coherence between test results and level in main subjects.

My conclusion could only be: RPM is not, in this Danish school, a good means of giving reliable information about the intelligence of the children tested. Quite apart from this, there is an interesting difference in the results from the first and fourth grades in this school. Why are the points scored by the first grades significantly higher than those of the fourth grades?

I wonder if an explanation may be approached from the following consideration: 80 to 90 per cent of the children in the first grades have attended kindergarten classes, and almost all the remainder have attended kindergarten right up to commencing school. They may thereby be assumed to be well stimulated with regard to cognition, and may possibly, as a result, have developed closer to their 'potential' than the English children, of the same ages, who were the criterion group in 1956. Using the same train of thought, if the fourth grade level lies relatively lower so that results are closer to the English 1956 standards, this could possibly be connected with the fact that English children of the same age, influenced by their schooling, had developed somewhat closer to their 'potential' than had the younger age groups.

Northwest Jutland. In comparison with Thailand there are very modest sociological-economic and other differences between large city and village in Denmark. We are in general a rich and highly developed homogeneous country. Certainly the northwest Jutland groups of children were very small, and chance may therefore play a major role. The results, all the same, tempt one to indicate that the test is not 'culturally fair' in our homogeneous country.

There are also, however, as can be seen from the concluding table (Table 1), big differences between the Copenhagen district school and northwest Jutland. Here again, there is a predominance of 'good' children, but it is amusing to ascertain that in northwest Jutland we find today a distribution of results that lie closer to a standard distribution than any of the other tested groups. In other words, distribution in northwest Jutland in 1979 is approximately like those of the English children who formed the basis of standardization in 1956.

In a village situated in a really thinly populated area with very poor agricultural soil one finds in 1979 a distribution of test results so good that it begins to make sense to believe that the test has operated as intended.

It is then something quite different to see if the results tell us anything about the one thing

<table>
<thead>
<tr>
<th>Grade</th>
<th>1st</th>
<th>4th</th>
<th>1st</th>
<th>4th</th>
<th>1st</th>
<th>4th</th>
<th>1st</th>
<th>4th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thailand</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Village</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17.0</td>
<td>11.8</td>
<td>21.3</td>
<td>7.8</td>
</tr>
<tr>
<td>Bangkok</td>
<td>0</td>
<td>16.0</td>
<td>6.7</td>
<td>29.0</td>
<td>56.7</td>
<td>29.0</td>
<td>30.0</td>
<td>28.0</td>
</tr>
<tr>
<td>Denmark</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Village</td>
<td>9.1</td>
<td>4.8</td>
<td>27.3</td>
<td>38.6</td>
<td>40.9</td>
<td>52.4</td>
<td>13.6</td>
<td>14.3</td>
</tr>
<tr>
<td>Copenhagen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>district</td>
<td>46.7</td>
<td>8.9</td>
<td>36.7</td>
<td>32.1</td>
<td>8.3</td>
<td>42.9</td>
<td>8.3</td>
<td>16.1</td>
</tr>
</tbody>
</table>

Results are given in percentages.
the test claims to measure: the general intelligence of the children. It does not appear to do so, as in these surroundings, too, there is only a modest agreement between the test results and the teachers' evaluations of the pupils' capabilities and level in school subjects.

To do justice, one must add: there is an agreement between teacher evaluation and test result for the fairly small number of pupils who come into the 'superior' group. This is probably, furthermore, an old experience in school psychological practice: a pupil may very well score low points in a Raven test and yet, with regard to school performance and intellect, function well, even especially well, whereas one has rarely seen a child score very high points in a Raven test without also in other ways demonstrating that s/he was well equipped.

Attempts at cautious conclusions

Finally, I quote the last lines of my original report for this little examination:

'As I have worked with applied school psychology for more than 20 years I have not much confidence in intelligence testing generally, and particularly not in the CPM (Coloured Progressive Matrices) test.

The results clearly show that CPM is neither "culture free" nor "culturally fair" towards the children tested both in Thailand and in Denmark.

My material is very limited, and investigations on a much larger scale are needed before general conclusions can possibly be made.

From my findings, I feel obliged to warn

Anders Poulsen
Dansk Psykologisk Forlag, Hans Knudsens Plads 1A, DK-2100 Copenhagen Ø, Denmark.'
intensively on a programme of known worth and has still not learnt as well as the other children in the group. He is by definition a slow learner and needs to be taught in a slower moving Distar group, either in normal or special schools.

One would like to conclude that, with this simple experimental design, namely one group with pre- and post-test measures, Distar Reading II caused these reading gains. Before one can do this one has to explain away other possible explanations. History, regression to the mean, maturation, testing or a combination of these may provide alternative explanations. Taking these in order: some event or sequence of events that happened at the school or at home between pre- and post-testing, whilst the programme was running, and not the programme itself, may have caused the reading gain. This can be refuted only by saying that no such events appear to have occurred in the teaching period. The reading gains are unlikely to be due to regression to the mean because of the selection procedure. Regression occurs when a group is chosen specifically because they have low scores on a reading test. Some of the children may have scored low because of chance events — an off day — so that, if they were retested at a later date after receiving no teaching programme, they would show reading gains. This is because they produced a score on the second testing that was closer to their true reading age. In this experiment the children were chosen using LEA screening data, they were given the Distar placement test and only then given the pre-test. Thus regression is unlikely to be the cause of the gains.

It may be that the children simply matured in the teaching time. However, the gains are so great and are in such contrast to previous slow progress that this seems unlikely.

The only real criticism is that the testers may have been biased in the post-test and been overly helpful or over-generous in scoring. However, since the tester was aware of this possible criticism, great efforts were made to be unbiased in the testing.

One may tentatively conclude that Distar did achieve these gains. The conclusion can become stronger the more times such findings as these are replicated with Distar Reading II.

If it has achieved anything, the experiment has shown how, with systematic teaching, seven children who were thought to be slow learners can achieve. This altered perception may work in the children's favour in future with their teachers.

Note: all the names have been changed to protect the pupils' confidentiality.

Bibliography


Dr Poulsen (School Psychology International, 4(3)), in common with many other critics of testing, fails to distinguish between the value of a test and the uses to which it is put. Paradoxically, it is only the availability of tests which allows psychologists to come to the conclusion that certain abilities, motivational dispositions or attitudes do not reflect stable individual differences but experience or the particular situation in which an individual has been placed (J. Raven, 1979).

Turning now to specifics, Dr Poulsen gives the impression that it is claimed both that the Progressive Matrices tests measure 'intelligence' and that they are 'culture free'. Yet nowhere in the Manual for the Progressive Matrices tests, or in the Guides which preceded it, is either of these claims made. On the contrary, attention is drawn to the extent to which test scores are influenced by fatigue and illness and to the importance of studying the environmental factors which influence performance. Attention is also drawn to the fact that Progressive Matrices scores correlate much less highly with school performance and scores on tests of general intelligence than do scores on verbal tests, such as the Mill Hill Vocabulary scale. (The latter, incidentally, can be administered even more quickly than the Matrices, taking only three or four minutes. If one's objective is to predict school performance, then this is the test to use.)

In relation to the claim that the test is not one of the best predictors of academic performance or general intelligence it is important to distinguish between Spearman's 'g', which Raven set out to measure, and general intelligence. It is also important to question Poulsen's assumption that measures of 'g' should not be subject to environmental influences.

Raven strongly opposed Burt's wish to define 'intelligence' as that part of general ability which is innate (J.C. Raven, 1956). On the contrary, he argued, 'intelligence', as the word is commonly used, is heavily dependent on acquired information (J.C. Raven, 1956, 1966). What he wanted to do was to measure its components separately, instead of having them inextricably jumbled up as they are in most measures of general intelligence. If that were done, he argued, 'intelligence' might turn out to be much more open to cultural influences than many people had suspected.

In this context, use of the Progressive Matrices tests has led to a number of findings which many psychologists have difficulty reconciling with their beliefs and assumptions. Thus the Progressive Matrices tests, scores on
which have so often been shown to be subject to major cultural and educational influences (see, e.g., Irvine, 1965; Silvey, 1972; Stallings and Kassovitz, 1974; Budhoff, Corman and Gimon, 1976), often scale equally well within different cultural groups (J. Raven 1981; Court, 1982; J.C. Raven, Court and Raven, 1980). It follows that the difference in average scores of different groups must be, in some sense, meaningful. They cannot be dismissed as irrelevant or 'absurd' simply because the mean scores of some groups are lower than expected. If the test works within the groups concerned one must investigate why it is that the scores of some groups were lower.

Again, clearly subject though scores are to cultural influences, there is remarkable similarity in the data obtained from different Westernized societies at any given moment over the past 40 years. Yet, equally, in all those societies there has been, during the past 15 years, a marked acceleration in development (J. Raven, 1981).

Likewise, although the Matrices tests are not particularly highly correlated with educational attainment and general intelligence, factor analytic studies regularly show that they load more highly than other tests on the 'g' factor. It would therefore seem that Raven lagged significantly in his objective of developing a test to measure 'g', but that, as he suspected, 'g' is not the stable disposition which many believe it to be. Contrary to what Burt believed, 'g' itself may therefore be open to social, cultural and educational influences as well as influenced by health and fatigue. If a test shows sensitivity to such variation it may therefore be that it is the assumption that that which is measured should be stable, and not the test in question, which is at fault. It may also be noted that the implication to be drawn from these facts is that analytic studies is that Raven was right to insist that 'general intelligence' is something which is heavily dependent on acquired information and, as such, different from 'g'.

Finally, in relation to Poulsen's observations about the 'peculiar' shape of the distribution of test scores (it is noted that J.C. Raven (1959) used the data he obtained with his own (and others') tests to question the notion that test scores 'should' be 'normally' distributed. He again argued that it might be the assumption about what reality should be like, and not the obtained distributions of test scores, which was wrong. For this reason he resisted the use of deviation IQs as the measure of expressing test scores. In fact, both the norms which Raven published and the graphs which have been published more recently (J.C. Raven, Court and Raven, 1977, with updates to 1983; J. Raven, 1981) reveal that the distributions which are obtained within age groups far from follow a normal curve.

In conclusion, therefore, while Dr Poulsen is right to alert test users to the dangers of drawing unjustified conclusions from the use of tests, he is wrong to imply that it is testing itself which is the problem. Dr Poulsen's real problem is that his critique does not go deep enough. Most of his assumptions need to be challenged. This applies particularly to the way in which he thinks about intelligence and ability. When people possess such a vast range of barely correlated abilities, motives and interests, which schools and society could help to develop and use, it is irresponsible, indeed immoral, for psychologists and will write in ways which allow teachers and managers to continue to believe that such constructs as 'intelligence' tap anything more than a fraction of the important variance in individual differences. It is our task to provide society with the concepts and measures required to implement manpower policies that are based on a framework that genuinely recognizes the importance of a very large number of independent talents, and involves values and judgements in addition to 'abilities'. Such complex motivational dispositions involve finely-tuned inter-dependence between cognitive, affective and behavioural capacities which cannot, therefore, be assessed separately. 'Level of cognitive development', or 'level of language development', cannot, therefore, be assessed meaningfully except in relation to a goal which the individual concerned cares very strongly about (J. Raven, 1960a, 1960b). The range of qualities which need to be assessed is discussed more fully in J.C. Raven (1956, 1966) and in J. Raven (1977a, 1977b). The implications for psychologists and educators is discussed in J. Raven (1982).

Bibliography


The 'State of States'

Reflections by USA State School Psychology Association Presidents

CAVEN S. MCLoughlin*, JOHN GUIDUBALDI* & JEANINE A. LIGHTEL†

*Kent State University and †Tallmadge City Schools, Kent, OH, USA

ABSTRACT The presidents of state school psychology associations affiliated to the National Association of School Psychologists were canvassed for reflections on five topics: (1) major challenges currently facing school psychologists in their state; (2) major challenges currently facing the state association; (3) state association accomplishments in the previous year; (4) state association priorities for the following year; and (5) state association actions anticipated or planned for the following year. Of the 43 state association presidents 38 responded (88 per cent). Categories of responses to each topic area are presented by state and region. Analyses of consensus responses within and across topic areas are given.

There has never been a time in the history of North American school psychology when the potential for collective action has been as visible. The profession's attempt to become proactive was amply illustrated by the 1981 Olympia school psychology leadership conference which was jointly sponsored by the National Association of School Psychologists (NASP) and the School Psychology Division of the American Psychology Association (APA-16). These two organizations consist of the primary professional affiliations for North American school psychologists. In attempting to project future conditions and to explore accommodating role changes, the invited leaders of the profession demonstrated at Olympia their optimism as well as their faith in collaborative endeavour.

This optimism appears justified in view of such professional advances as marked expansion in national organization memberships, rapidly increasing numbers of state associations, dramatic growth of training programmes and resulting increases in school psychology graduates, and more rigorous certification standards. Additionally, the capabilities of national and state-level organizations have increased to include more effective vehicles for communication, higher quality in-service training opportunities and more influential legislative lobbying activities. More detailed explanation of the historical and current trends in school psychology in the USA has been presented elsewhere (Tindall, 1979; Fagan, 1981; Guidubaldi, 1981).

A recent survey of school psychology state association presidents by the NASP Assistance to States Committee confirms the growing national cohesion of the school psychology profession and reflects an increasing consensus regarding priorities, concerns and practices. On behalf of their associations, presidents were specifically asked to address the following five topics: (1) major challenges currently facing school psychologists in their state; (2) major challenges currently facing the state association; (3) state association accomplishments in 1980-1981; (4) state association priorities for 1981 to 1982; and (5)

John Raven  Scottish Council for Research in Education, 15 St John Street, Edinburgh EH8 6JR, UK.