



# NSW COUNSELLORS' BULLETIN ACER MEASUREMENT UNIT

27 FEB 1973

No. 2 Australian Council for Educational Research  
Hawthorn, Victoria

July 1972

## INTRODUCTION

In this issue a description is presented of an ongoing study of the ACER Test of Learning Ability (TOLA). It was decided by ACER early last year to carry out a series of studies investigating a number of aspects of the test, such as its validity and the effect of reading ability on test scores.

The first study (essentially a pilot study) was carried out in October last year (1971) and was primarily an investigation of the concurrent validity of the test. In August this year a further study will be carried out, examining the effect of the verbal nature of the test on the measurement of general ability and consequently on the assessment of learning potential.

### A. The October 1971 Study

The study was carried out in two Melbourne schools with 124-128 Grade 4 children (4 classes) and 130 Grade 6 children (4 classes). The tests were administered on two consecutive mornings, the order of the presentation being reversed for 2 classes in each sample.

#### 1. The tests were as follows:

##### (a) Tests administered to both Grade 4 and Grade 6:

- (i) Primary Mental Abilities - Number Facility. (2 parts)  
'The ability to work with numbers, to handle simple quantitative problems rapidly and accurately, and to understand and recognize quantitative differences.'  
Part one was a number sense test involving series and problems such as 'Donald bought two candy bars that cost five cents each. How much did he spend altogether?'  
Part two was a test of addition, such tests being regarded as 'pure' measures of number ability.
- (ii) Primary Mental Abilities - Verbal Meaning. (2 parts)  
'The ability to understand ideas expressed in words.'  
Part one was a straight forward vocabulary test.  
Part two was similar, but presentation of the stimulus word was oral, the children having to select the corresponding picture.
- (iii) Primary Mental Abilities - Reasoning. (2 parts)  
'The ability to solve logical problems.'  
Part one was a figure grouping test in which children selected the figure unlike those in the remainder of a set.  
Part two was a word grouping test in which children selected the word (or concept) which did not belong with the other words (or concepts) in the set.
- (iv) The OTIS AB wide range traditional spiral omnibus test.

- (v) PEP Mathematics Test M3. Operations on counting numbers.
  - (vi) PEP Mathematics Test M5. Problems.
  - (vii) ACER Mathematics Test AM4. Whole Numbers Part IV Computation.
- (b) Tests administered to Grade 4 only:
- ACER Test of Learning Ability TOLA 4 (Parts 1, Vocabulary; 2, Numerical; and 3, Reasoning).
- (c) Tests administered to Grade 6 only:
- (i) ACER Intermediate Test E.
  - (ii) Trial form of ACER Test of Learning Ability TOLA 6 (Parts 1, 2, and 3 as for TOLA 4).

## 2. Results

Means, and Standard Deviations were calculated (including reliability estimates where appropriate). These are shown in Table 1, page 3. (PMA results for Grade 6 were not calculated for the separate parts of each subtest.) Generally the tests show a negative skew. This is partly due to the nature of the sample which was above average as may be seen from the data in Table 2 (NSW data was only available for total scores).

TABLE 2: Comparisons of TOLA 4 Means and Standard Deviations

SAMPLE	TEST 1		TEST 2		TEST 3		TOTAL	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Oct.71 Study	20.008	5.306	12.744	3.279	12.480	3.019	45.232	9.722
National Norm	14.1	5.9	9.8	3.1	10.6	3.6	34.6	10.9
NSW Norming	-	-	-	-	-	-	37.430	12.616

The children in the sample were superior to those tested in the NSW Norming Study (May-June 1971) and the ACER National Norming Study (April-May 1971). The group consisted of children selected from upper socio-economic areas\* and were expected to be fairly even in ability thus reducing the variance of the sample. This, in effect, depressed both the correlations and the reliabilities. [Coefficient alpha was calculated for the three parts and the total of TOLA 4 in the ACER National Norming, obtained values being 0.83, 0.61, 0.71 and 0.89 respectively. Estimating KR20 by Tucker's approximation for the NSW sample (the most heterogeneous of the three samples) gave a reliability of approximately 0.91.]

\* Since the sample was small, classes from these areas were selected to avoid problems of children with language problems (migrants) which would reduce the sample size further.

TABLE 1: Means, Standard Deviations, and Estimated Reliabilities\*

TEST	NUMBER TESTED		MEAN		POSSIBLE SCORE	STANDARD DEVIATION		EST RELIABILITY	
	Gr.4	Gr.6	Gr.4	Gr.6		Gr.4	Gr.6	Gr.4	Gr.6
OTIS AB	128	130	37.117	54.792	69	10.961	9.743	.886	.916
PMA NUMBER 1	128		15.203		20	3.417		.775	
PMA NUMBER 2	128		12.258		30	4.062			
PMA NUMBER TOTAL	128	130	27.445	35.162	50	6.250	6.103	.733	.772
PMA REASONING 1	127		16.976		25	4.240		.767	
PMA REASONING 2	127		16.638		25	3.906		.710	
PMA REASONING TOTAL	127	130	33.614	39.323	50	7.174	5.149	.829	.751
PMA VERBAL 1	125		18.080		30	5.221		.793	
PMA VERBAL 2	125		18.200		30	4.122		.649	
PMA VERBAL TOTAL	125	130	36.280	49.208	60	8.303	6.907	.830	.864
PMA TOTAL	125		97.896		160	17.376		.894	
PEP M3	124	130	18.210	30.292	60	6.457	8.640	.749	.836
PEP M5	125	130	15.520	23.434	60	4.978	5.833	.614	.641
AM4 PT IV	127	130	22.906	40.577	49	7.293	5.933	.813	.859
TOLA 4 1	125		20.008		31	5.306		.805	
TOLA 4 2	125		12.744		22	3.279		.586	
TOLA 4 3	125		12.480		18	3.019		.672	
TOLA 4 TOTAL	125		45.232		71	9.722		.859	
TOLA 6 1		130		28.192	44		6.936		.834
TOLA 6 2		130		13.446	24		4.074		.713
TOLA 6 3		130		16.285	22		3.807		.787
TOLA 6 TOTAL		130		57.923	90		12.287		.890
INTER E		130		44.969	75		13.141		.921

\* Reliabilities estimated by Tucker's (1949) approximation to KR-20.

Intercorrelations were calculated between the tests for both grade levels, and were examined in four sections, Verbal, Numerical, Reasoning and Total Test Scores.

(i) Verbal Abilities

(a) Grade 4 data

TOLA 4 Vocabulary section correlated with PMA Verbal Meaning .721, with the PMA Vocabulary subsection .688 and with the Picture Vocabulary subsection .581. This indicated that the first part of TOLA 4 was measuring approximately the same abilities as the PMA Verbal Meaning Test.

(b) Grade 6 data

TOLA 6 Vocabulary section correlated with PMA Verbal Meaning .685. This was somewhat lower than the Grade 4 correlation and was mainly due to the extensive negative skew on the PMA distribution. (This may be seen in examination of Table 1 data.) Because of the fact that TOLA 6 was essentially a trial form, and because of the distribution skew mentioned earlier (most marked for the Picture Vocabulary section), no correlations were calculated for the subsections of the PMA Verbal Meaning section.

(ii) Numerical Ability

(a) Grade 4 data

The intercorrelations for the Numerical Ability tests for Grade 4 are shown in Table 3.

TABLE 3: Grade 4 Intercorrelations of Tests Designed to Measure Numerical Abilities

	PMA (NA)	(Opns) M3	(Comp) M5	(Problems) AM4
TOLA 4 (2)	.518	.451	.452	.544
PMA (NA)		.533	.491	.652
M3			.627	.625
M5				.609

It can be seen that TOLA 4 (Part 2 - Numerical) correlated positively and significantly with AM4 and PMA (Numerical Ability), although neither of these correlations was particularly high. This is due partly to the fact that these tests have relatively few items. If it was thought desirable to study further the relationship between learning potential and numerical ability then a longer test of the AM4 type would be suitable at this level.

Correlations between TOLA 4 (Part 2 - Numerical) and PMA Problems and Addition separately were .526 and .359. The latter test was extremely speeded and possibly accounts for the low correlation.

(b) Grade 6 data

The intercorrelations for the numerical ability tests at Grade 6 are shown in Table 4.

TABLE 4: Grade 6 Intercorrelations of Tests Designed to Measure Numerical Abilities

	PMA (NA)	(Opns) M3	(Comp) M5	(Problems) AM4
(Trial) TOLA 6 (2)	.438	.645	.665	.473
PMA (NA)		.628	.540	.557
M3			.833	.676
M5				.653

At the Grade 6 level TOLA 6 correlated better with the M series tests, although again correlations were not high. AM4 was a less useful test at this level being too easy as may be seen from the information in Table 1. TOLA 6 correlated fairly poorly with the PMA Numerical Ability tests, probably because the latter were too easy.

(iii) Reasoning Ability

(a) Grade 4 data

TOLA 4 (Part 3 - Reasoning) correlated with the PMA Reasoning Test .471, which was to be expected in view of the fact that neither parts of the PMA Reasoning section were very similar to the comparable section in TOLA. Correlation with the first part (the non-verbal figure grouping) was lower, being .397.

(b) Grade 6 data

TOLA 6 (Part 3 - Reasoning) correlated .471 with the PMA Reasoning test. This was exactly the same result as was obtained with the Grade 4 data.

(iv) Total Test

Although TOLA appears to be separated into subtests measuring separate abilities, the subtests exist only for ease of administration of different item types and the test was designed primarily as a single score test and consequently the total scores are the most relevant data.

(a) Grade 4 data

TOLA 4 correlated with the PMA total .812 and with OTIS AB .834. PMA and OTIS AB correlated .781. This provided

substantial evidence for the concurrent validity of TOLA 4 at Grade 4.

(b) Grade 6 data

The correlations among TOLA 6, Total PMA, OTIS AB and ACER Intermediate E are shown in Table 5.

TABLE 5: Grade 6 Intercorrelations among Tests Designed to Measure General Intellectual Ability

	Total PMA	OTIS AB	Inter.E.
TOLA 6	.749	.828	.823
Total PMA		.789	.799
OTIS AB			.838

As for the Grade 4 data TOLA correlates better with the omnibus tests OTIS AB and Intermediate E than with the multi-aptitude PMA. This is not surprising.

In summary it appeared from examination of the correlations that TOLA has good concurrent validity, both with respect to the three item types and the test as a whole.

### 3. Further Analyses of the data

In addition, further analyses were performed on the data to verify the conclusions reached.

The two matrices of correlations were submitted to a Principal Axis analysis with Varimax rotation. Communalities were estimated by using highest row values, and the criterion for the extraction of factors was Eigenvalues greater than 0.95. For both sets of data only one factor was extracted (Intermediate E was unfortunately omitted from the analysis of 6th grade data). Factor loadings are shown in Table 6 (page 7).

From this analysis it is a reasonable assumption that TOLA is a useful measure of general intellectual ability.

To further examine the relations among the subtest, looking for differences among the tests in particular not detected by the principal axis analysis, a Guttman-Lingoes Smallest Space Analysis (SSA-1) was performed on the correlations. This technique enables tests to be positioned in n-dimensional space (in this case 2 dimensional), such that the correlations between tests are represented by the distances between the points assigned to them in this space. Tests which are closely correlated may be identified by partitioning the space into domains. The results are shown in Figures 1 and 2 on pages 8 and 9.

TABLE 6: Factor Loadings on Tests at Grade 4 and Grade 6

Test	Grade 4 Loading	Grade 6 Loading
TOLA 4 or 6 (Pt 1 - Vocabulary)	0.812	0.729
TOLA 4 or 6 (Pt 2 - Numerical)	0.744	0.765
TOLA 4 or 6 (Pt 3 - Reasoning)	0.748	0.803
Total 4 or 6 (Total)	<u>0.933</u>	<u>0.923</u>
PMA (Pt 1 - Numerical)	0.778	0.632
PMA (Pt 2 - Reasoning)	0.648	0.595
PMA (Pt 3 - Verbal)	0.815	0.786
PMA (Total)	<u>0.925</u>	not included
AM4	0.762	0.688
M3	0.672	0.851
M5	0.696	0.863
OTIS AB	<u>0.888</u>	<u>0.897</u>

It can be seen that TOLA total scores fall into an area in which Verbal, Reasoning and Numerical Domains overlap, indicating that the test as a whole sampled the necessary skills. The data would appear to support a multi-factor approach to intelligence with Verbal, Reasoning and Numerical Domains. (However it must be pointed out that these results should be interpreted with some caution as the significance tests for the technique just failed to reach a satisfactory level for a two dimensional solution. Representation in three dimensions may have given a better description of the strengths of correlations between tests, but this would be harder to illustrate.)

The above analyses provided representations of the relationships among the tests which could be said to correspond to two models of intelligence - the general ability model (as illustrated by the Principal Axis Analysis) and the multi-abilities model (illustrated by the Smallest Space Analysis). The data were then further analysed by the technique of hierarchical cluster analysis to see if results could be obtained which would, to some extent, reconcile these two models. The results are shown in Figures 3 and 4.

Hierarchical cluster analysis is a technique for analysing the relationships among a group of tests as they lie in n-dimensional space (in this case three dimensional) as determined by a particular factor analytic technique. The cluster analysis selects the pair of tests (among, in the Grade 4 case, the 11 tests) which are closest in the space. Then it selects the two next most similar among the remaining 'clusters' (in the Grade 4 case nine separate test 'clusters')

FIG.1 SSA-1 plot of Grade 4 tests in two dimensions

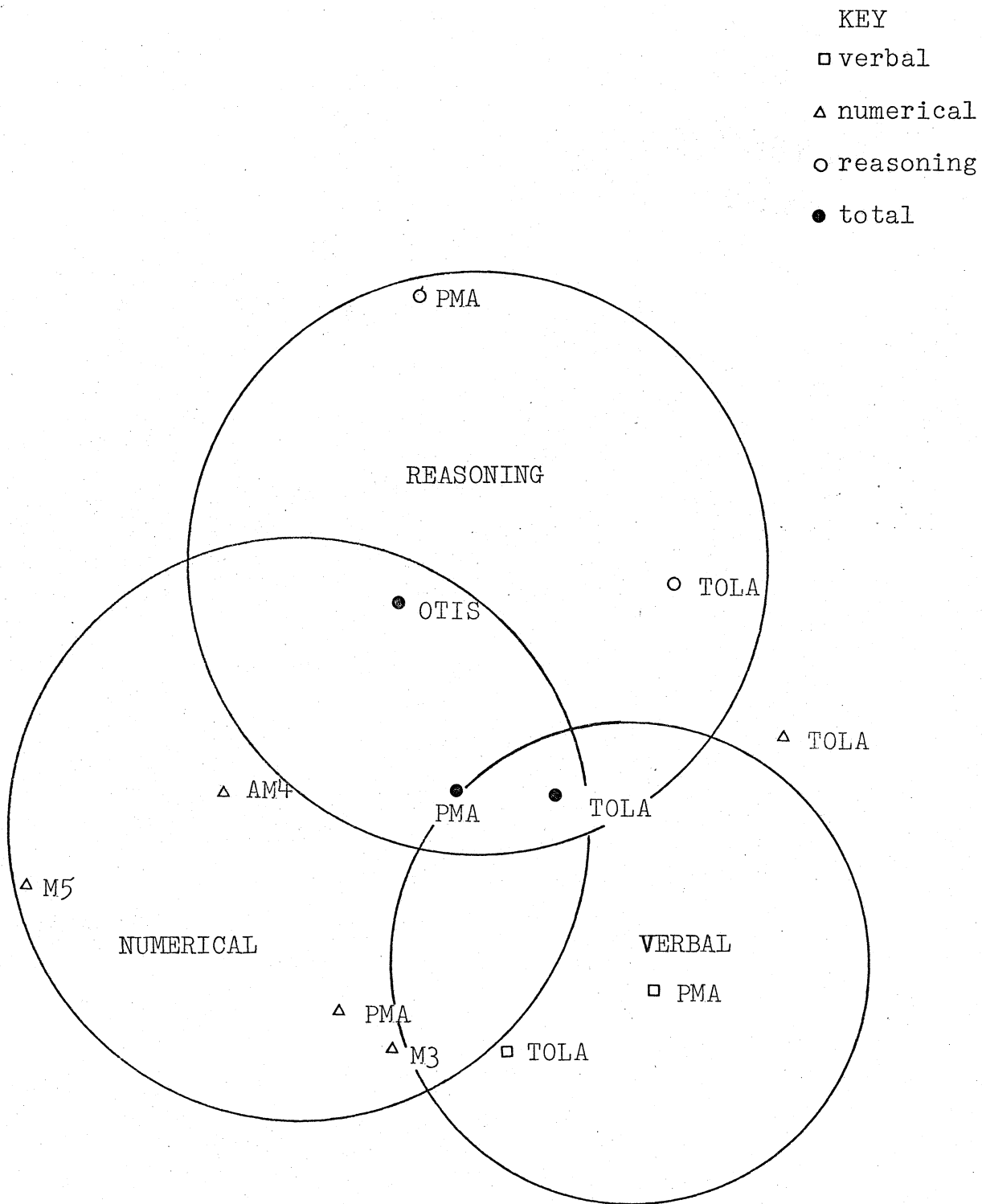
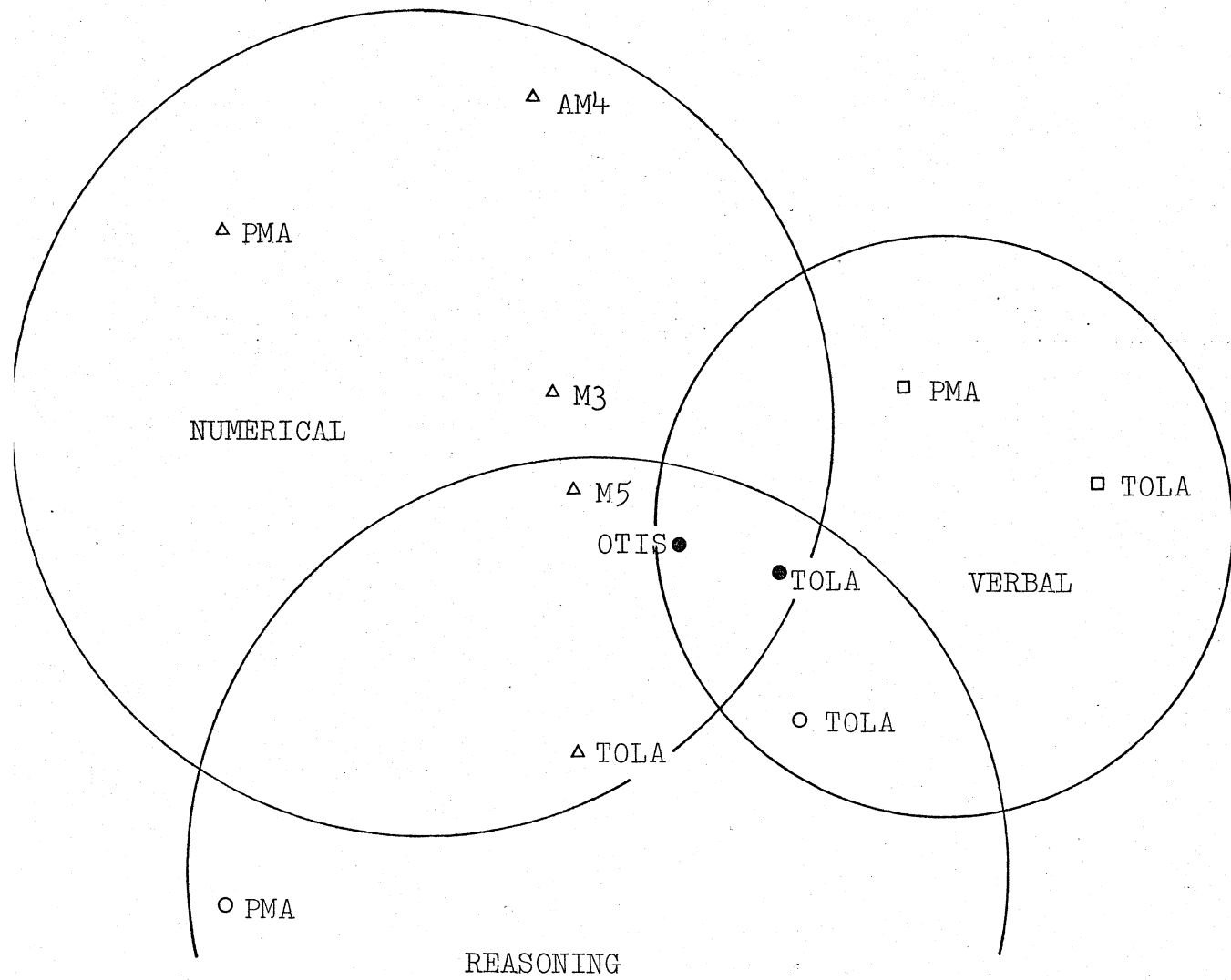




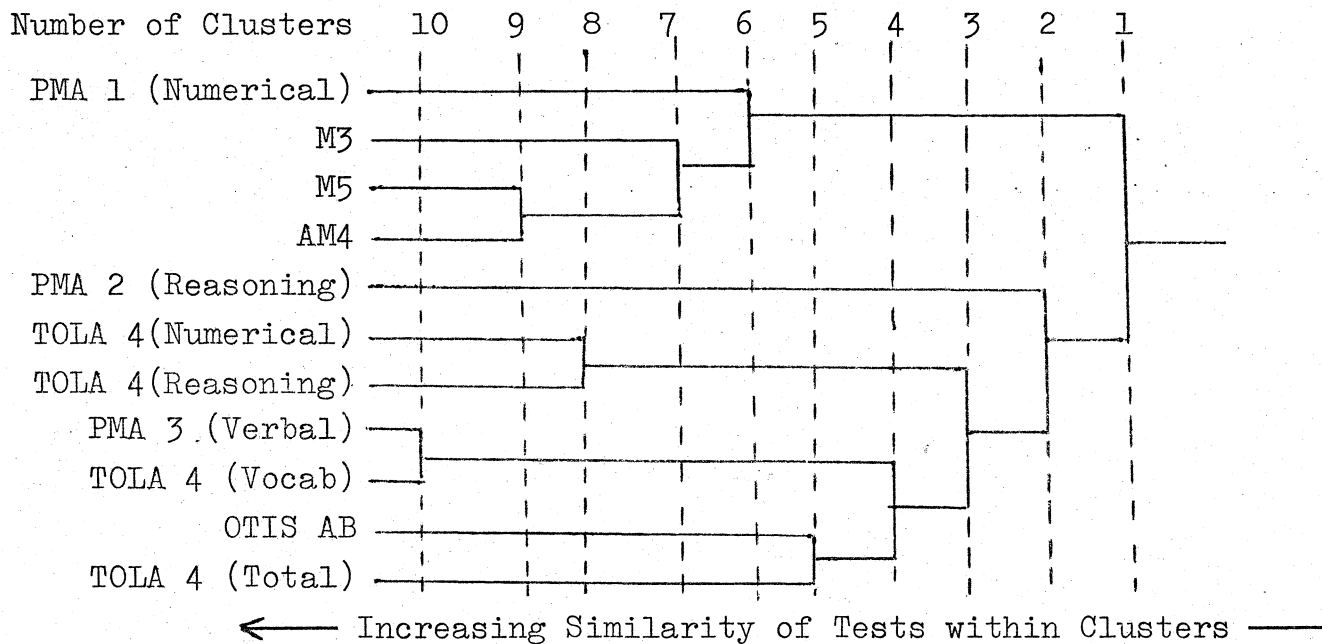
FIG.2 SSA-1 plot of Grade 6 tests in two dimensions

KEY  
 □ verbal  
 △ numerical  
 ○ reasoning  
 ● total



and one two-test 'cluster') and so on until only one cluster is left (with all tests in) it can be thus seen to build up a hierarchy of clusters according to similarity. Hence in Figure 3, the cluster analysis of the Grade 4 data, the two tests most similar were TOLA 4 (Vocab) and PMA 3 (Verbal). The next most similar pair were M5 and AM4, and the next two most similar TOLA 4 (Reasoning) and TOLA 4 (Numerical). The next most similar pair was formed by M3 and the M5-AM4 cluster. This pairing procedure continued until only one pair was left to be formed.

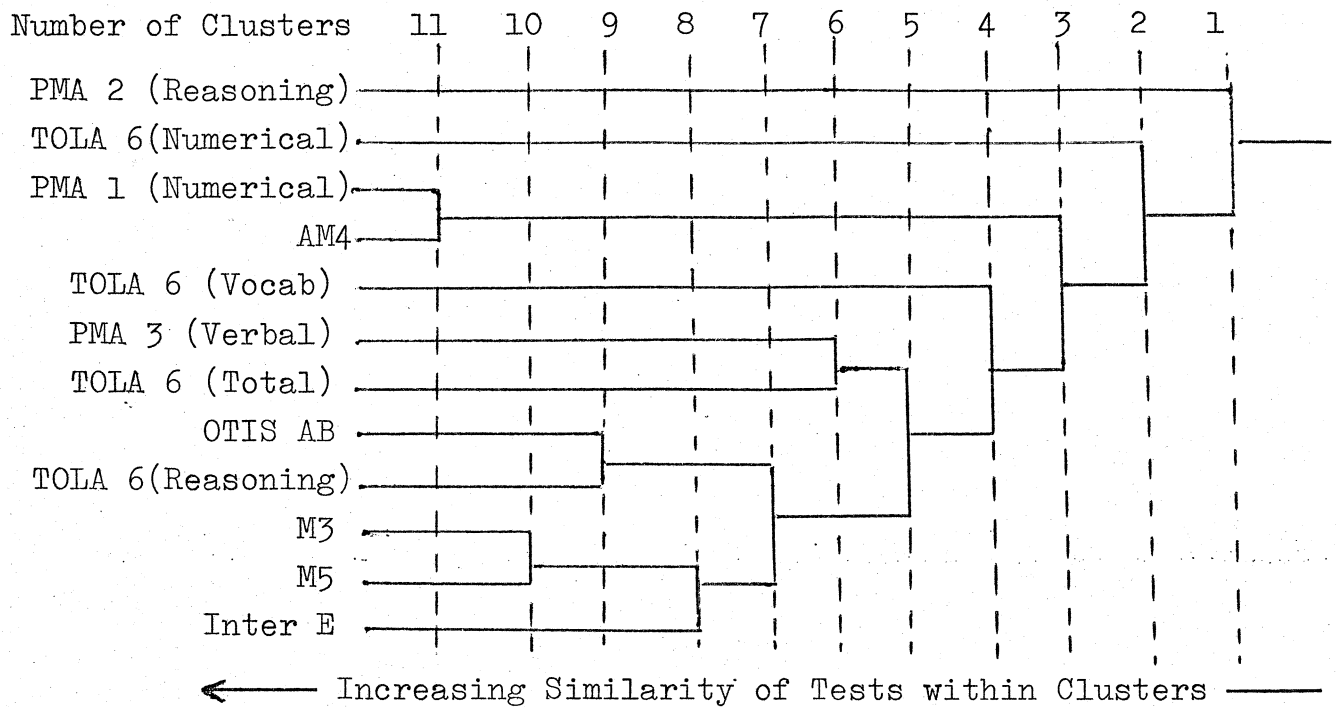
FIGURE 3. Hierarchical Cluster Analysis of Grade 4 Data



For the Grade 4 data two basic hierarchies were formed. One was a numerical clustering with M5 and AM4 being most similar and M3 and PMA 1 (Numerical) also in the cluster. TOLA 4 (Numerical) did not fall into this cluster but grouped instead with TOLA (Reasoning). This latter cluster was linked with the verbal cluster [PMA 3 (Verbal) and TOLA 4 (Vocabulary)] and the total test cluster (Total TOLA 4 and OTIS AB. The PMA 2 Reasoning test did not cluster with the other tests.

At the Grade 6 level the picture was less definite (see Figure 4). Two separate numerical clusters formed: PMA 1 (Numerical) - AM4, and M3-M5. This latter cluster (M3-M5) was then amalgamated with other clusters of total test scores and TOLA 6 Reasoning. TOLA 6 (Vocab) was slightly similar to PMA 3 (Verbal) and the total tests, while TOLA 6 (Numerical) was unlike most other tests. PMA 2 (Reasoning) was still the most independent, however.

FIGURE 4. Hierarchical Cluster Analysis of Grade 6 Data



This analysis gave some idea of how different abilities may be measured by the same test at different levels of age or grade.

The evidence suggests that TOLA is a valid measure of general intelligence, although a significant relationship could be seen (in the hierarchical cluster analysis and smallest space analysis) between verbal measures and TOLA total test scores. It was thus concluded that some attempt should be made to examine the effect of reading ability on the measures of general intelligence provided by verbal and non-verbal tests.

B. The August 1972 Study

Early in August it is proposed to carry out a study to investigate the issue raised by the October 1971 study and mentioned in the previous paragraph. Chiefly it is proposed to examine the relationship between reading ability and TOLA scores (at both Grade 4 and Grade 6 levels) compared with other verbal and non-verbal tests of general ability. The tests used will be PAT reading tests, TOLA 4, TOLA 6, and the Lorge Thorndike Battery (with verbal and non-verbal components). The TOLA 4 data will be supplied from the Grade 4 testing this year and the trial testing of some replacement items. The TOLA 6 test will be comprised of performance on two trial forms each with Vocabulary, Mathematics and Reasoning sections as in TOLA 4.

In addition it is hoped to obtain a further criterion measure, (against which TOLA may be evaluated) from teachers' global assessments of their pupils' learning potential or ability.

This testing will be carried out in ten Sydney metropolitan schools by members of the Test Advisory Committee assisted by a number of counsellors in training. Analysis of the data and preparation of a report will be carried out by ACER.