New Zealand norms for the Visual Sequential Memory Subtest of the Illinois Test of Psycholinguistic Abilities - Revised

Sandra Heriot

Ivan Beale

University of Auckland

The Visual Sequential Memory (VSM) subtest of the Illinois Test of Psycholinguistic Abilities - Revised often is used in the neuropsychological assessment of learning disabilities. However, data on the validity of the diagnostic use of this test are very limited. The aim of this study was to extend earlier work by testing larger samples of children and thereby provide comparative data on the applicability of USA norms to New Zealand conditions. 238 participants aged from 6 to 16 years were tested on the VSM and two modified versions of the test, one using random shapes and the other using common objects. While the performance of New Zealand 7- to 10-year-olds was similar to US children of the same ages, performance by New Zealand 6-year-old children was significantly better than US 6-year-olds. Therefore, there is reason to prefer the norms developed in this study over the existing norms for US children.

In 1961, the early form of the Illinois Test of Psycholinguistic Abilities (ITPA) was developed and published in an experimental edition (Kirk & McCarthy, 1961). The revised edition (Kirk, McCarthy, & Kirk, 1968), like the original ITPA, was conceived as a diagnostic, rather than as a classificatory, test of psycholinguistic abilities.

According to Kirk and McCarthy (1961), the clinical model of the ITPA is based on three dimensions of language function: a) channels of communication (auditory input, vocal output and visual input, motor output); b) levels of organisation (representational, automatic); and, c) psycholinguistic processes (receptive, expressive, organising).

The revised edition of the ITPA was standardised on 962 children, aged 2 - 10 years, each of whom met 5 specific criteria: a) average intellectual functioning; b) average school achievement; c) sensory-motor integrity; d) at least average characteristics of personal-social adjustment; and, e) English spoken as the family language. The children were drawn from five medium-sized cities located in the Midwest of the USA. The ITPA-R is recommended for use with children aged from 4 - 8-years-old, however, norms are provided for the range from 2 - 10 years to permit interpretation of results for children scoring below or above the 4 - 8 year norms (Preston, 1984, pp 354-364).

The use of the ITPA-R with language-disabled children has indicated that the most potent diagnostic subtests are at the automatic level, especially those for sequential memory (Zaidel, 1979). The Visual Sequential Memory (VSM) is one test represented at this level; it involves the visual-motor channel and organising process. For example, a number of studies have noted the diagnostic and predictive usefulness of the VSM task in the assessment of reading disabilities (Amoriell, 1979; Hirshoren, 1969; Kass, 1966; Pavlidis & Fisher, 1986). The VSM subtest assesses the ability to reproduce sequences of abstract figures from memory. In this test, a card is presented for five seconds, displaying a sequence of between two and eight figures. The card is removed and the participant is required to replicate the sequence using individual plastic tiles.

We gratefully acknowledge advice and assistance from Dr I.Field and M.Henning and thank the schools that participated in this study.

There are 25 items and 2 trials can be given per item. Clearly, performance on such a task could be facilitated or circumvented by using mnemonic strategies or by verbal codes so as to utilise auditory memory; the use of abstract figures is intended to counteract this tendency (Kirk & Kirk, 1971).

A few studies have examined the appropriateness of using US norms in other countries. Mittler and Ward (1970) found that the scores of four-year-old British children were similar to those of the US standardisation sample of the same age. Teasdale and Wray (1975) used similar selection criteria to those used for the ITPA-R norms and found no differences in raw scores between Australian children aged 5:7 years to 6:1 years and those in the US sample. However, the US children aged 6:7 years to 7:1 years were significantly better than Australian children of the same age. St George (1972) reported that mean raw scores for NZ children aged five and six years corresponded to US norms for less than five years. Gronwall (1980) sampled small groups of children within the ITPA-R age range to check whether the US norms for the VSM subtest were appropriate for use in NZ. She also examined performance in older children and young adults to see if an extension of its use beyond the normed age groups would be justified. No ceiling effects were found in her study and she concluded that the US norms for the VSM subtest may be adequate for Auckland children whose ages fall within the ITPA-R reference group. However, because the variance among the NZ sample was greater than in the US sample, she suggested that a more conservative comparison between scores be used than that advocated in the ITPA-R manual.

A child's score on a norm-referenced test reflects his or her performance relative to the performance of children on whom the test was standardised. It is an essential practice for users of normative tests to interpret an obtained score with reference to sets of norms appropriate for the individual tested and for the intended use (A.P.A., 1974). At present, there is little standardisation data available for educational and psychological tests used in NZ. Essentially, the reliability and validity of these tests for use with NZ children is unknown (Ballard, 1988). Despite the widespread use of the ITPA-R (Salvia & Ysseldyke, 1991) and on-going use of the VSM test in clinical settings (Wilson, 1992), no data on the validity of the VSM test for NZ use have been gathered since the early 1980s. One aim of this study, therefore, was to extend the earlier work of Gronwall (1980) by testing larger samples of children and thereby providing more information on the applicability of the US norms to the NZ population.

This paper reports part of a larger study which also examined the validity of the VSM subtest as a measure of visual sequential memory. The VSM subtest was compared with two modified versions of this test - one version used stimulus items that were readily verbally coded (common objects) and the other version used stimulus items of low verbal codability (random shapes). Procedures and results from the larger study are reported in this paper where relevant.

METHOD

Participants

There were 238 school students (136 female, 102 male) who participated in this study. The age range was from 6 - 16 years. Participants were divided into 11 age groups based on their birthdates. Participant information is summarised in Table 1.

TABLE 1. *Participant characteristics (number, mean age, standard deviation, sex distribution) for each age group.*

Age Group (years)	Number	Mean age (months)	Standard Deviation	Female (n)	Male (n)
6	28	78.71	2.94	17	11
7	26	89.73	2.86	11	15
8	23	101.22	3.19	12	11
9	22	114.82	3.16	14	8
10	25	126.28	3.41	15	10
11	22	138.23	3.52	12	10
12	25	151.72	2.91	13	12
13	14	162.21	3.14	9	5
14	23	175.13	3.32	13	10
15	13	183.69	2.96	8	5
16	17	197.77	3.15	12	5

Five schools were chosen for this study: one coeducational primary school; one coeducational intermediate school; one boys' and one girls' secondary school; and, one coeducational secondary school.

Schools were located in areas that represent average socioeconomic distribution for the Auckland region (Crothers, 1992). No exclusionary criteria were set for this study. Each school was asked to provide students from classes that represented mixed levels of educational ability. Participant selection was then based on the teacher's consent and the student's willingness to participate in the study. Prior to the study, an information sheet was given to each participant and written consent to participation was obtained from their parent or guardian.

Stimulus Materials

The standard VSM materials were used: display booklet; clock with second-hand; tile board; and, a set of 17 individual white plastic tiles (26cm x 26cm each) which displayed an arbitrary visual symbol. The figures used in the VSM test were engraved in black onto the tiles. Two other versions of the VSM, using different sets of symbols, also were used. These are referred to as Random Shapes (RS) and Common objects (CO).

Procedure

All testing was carried out by the examiner in a quiet room, at the participant's school. Total testing in each session was approximately 30 - 40 minutes' duration for each participant, approximately 10 - 15 minutes each for the VSM, RS and CO subtests. Each test was administered separately, with the order of tests randomised for each age group, so as to minimise possible position effects. The tests were presented to each participant in one of the following orders and recorded for each participant:

Order	1: VSM,	RS,	CO
Order	2: VSM,	CO,	RS
Order	3: RS,	VSM,	CO
Order	4: CO,	VSM,	RS
Order	5 : RS,	CO,	VSM
Order	6: CO,	RS,	VSM

The administration of all three tasks was in accordance with the instructions in the ITPA-R Manual (Kirk, McCarthy, & Kirk, 1968) for the VSM subtest. The general administration procedure involves a demonstration trial (DEMO II, for children aged six years and older), followed by a sampling procedure, after which basal and ceiling levels are obtained. Only those tiles used for each test sequence were available to the participant. Each participant was given a full demonstration in which the examiner performed the task first. A sequence card in the test booklet was exposed to the participant for five seconds. The sequence card was removed and the sequence was replicated by the examiner, by placing corresponding tiles on the tray in the same order. The participant was then asked to perform the task. The sequence card was reexposed and feedback on the task was given to the participant.

After the demonstration, selected items were sampled until the participant failed on the first trial. When a failure occurred, the participant was told that it was "not quite right". No other type of feedback about performance was given to the participant for any test items. The participant was presented with a second trial and sampling was discontinued. The test presentation then proceeded backwards through unadministered items until three consecutively listed (although not necessarily consecutively presented) test items were passed on the first trial. The examiner then returned to the point where the first failure occurred in sampling and continued testing until two consecutively listed test items were failed on both trials, and beyond any successes achieved in sampling.

The participant was given two points for each item passed on the first trial, one point for each item passed on the second trial and no points for items failed on both trials. In addition, two points were given for each item below the basal level. Orientation of the tiles was disregarded in scoring. All record sheets were checked by an independent rater for accuracy in scoring basal and ceiling levels and overall test scores. Agreement between the examiner and the rater was 100% for scoring basal and ceiling levels and 99.99% for overall test scores. Test scores were corrected prior to data entry. At the end of the testing session, participants were given a reward (e.g., sticker, chocolate bar) for their participation in this study.

RESULTS

The results of this study are presented and analysed in the following ways. First, in Section 1, data are presented from the earlier work by Gronwall (1980) and from the original standardisation sample of Paraskevopoulos and Kirk (1969). Data from the VSM subtest used in this study also are presented. Section 2 focuses on comparisons of the original US VSM data with the NZ VSM data for corresponding age groups (i.e., 6- to 10-year-old groups). Finally in Section 3, age, gender and test order variables are examined using analysis of variance.

In Sections 2 and 3, least-squares means (lsm) were used because of the unbalanced samples in this study and to control for the effects of the independent variables. Where main effects were found in the analyses of variance, Bonferroni-t protected post hoc comparisons were carried out on the least-squares means. Bonferroni-t was used to control for Experimentwise Type I errors. The critical level was adjusted for the number of possible combinations. This provided a more conservative level of significance required to reject the null hypothesis. Because the level of significance varied depending on the number of comparisons, it has been reported for each analysis.

In Section 3, age groups have been collapsed into

five age categories: Category 1 = 6 and 7 years; Category 2 = 8 and 9 years; Category 3 = 10 and 11 years; Category 4 = 12 and 13 years; and, Category 5 = 14, 15 and 16 years. This is justified by there being no significant differences between means for the ages collapsed into each category.

1. Presentation of Data

Data from Gronwall (1980) that relate to the age groups used in this study are presented in Table 2. Although these data will not be analysed further, they have been presented here to illustrate existing data on the performance of NZ children on the VSM subtest.

The original standardisation sample data, obtained from Paraskevopoulos and Kirk (1969), were converted to mean raw scores, standard deviations and ranges. These are presented in Table 3. There was an increase in scores over the age groups and the variance in scores was reasonably low within each age group and within a similar range across groups.

TABLE 2. *Mean raw scores, standard deviations and ranges on the standard VSM subtest from Gronwall (1980) data*

Age group (years)	Mean Age (years)	Number	Mean Raw Score	Standard Deviation	Range
6 - 8	7.17	15	19.27	4.32	7-27
10 - 12	11.00	15	24.13	4.96	19-36
13	13.64	15	23.67	4.95	16-31
15	15.60	15	29.20	5.85	18-39

TABLE 3. *Mean raw scores, standard deviations and ranges for each age groupon the original VSM subtest from Paraskevopoulos and Kirk (1969) data.*

Age Group (years)	Number	Mean Raw Score	Standard Deviation	Range
6	130	17.89	3.36	4-27
7	107	20.02	3.67	12-31
8	132	21.24	3.70	14-30
9	130	22.71	4.23	14-37
10	40	23.05	3.31	17-30

The data obtained from the present study were converted to mean raw scores, standard deviations and ranges for each age group on the VSM subtest. These are presented in Table 4. The overall trend on the VSM was for mean raw scores to increase, although somewhat unevenly, across age groups.

TABLE 4. *Mean raw scores, standard deviations and ranges for each age group on the standard VSM subtest*

Age Group (years)	Mean Raw Score	Standard Deviation	Range
6	20.82	4.30	15 - 36
7	20.56	3.43	15 - 28
8	20.48	2.50	16 - 24
9	23.77	3.70	18 - 29
10	25.20	5.51	20 - 43
11	27.36	5.72	18 - 38
12	25.44	4.65	20 - 35
13	27.00	6.47	16 - 38
14	27.78	6.18	19 - 48
15	30.46	7.21	22 - 42
16	30.82	7.08	21 - 44

2. Comparison of NZ and US VSM Data

Analysis of variance was used to compare NZ and US VSM scores on age (6 - 10 year olds), gender and group (US versus NZ) variables. For the combined data set, a significant effect was found for age, $F(4,642)=18.71$, $p=.0001$. Post-hoc comparisons for means of 6-, 7-, and 8-year-olds were significantly lower than means for 9- and 10-year-olds, $p<.005$. Other age-group comparisons were nonsignificant. A significant group effect was found, $F(1,642)=11.14$, $p<.001$. The difference between means for the NZ group (lsm=22.24) and for the US group (lsm=20.93) was significant, $p<.001$.

A significant interaction between group and age also was found, $F(4,642) = 3.62$, $p<.01$. Post-hoc comparisons showed that the mean for NZ 6-year-olds (lsm=21.17) was significantly larger than US 6-year-olds (lsm=17.86), $p<.0011$. No other significant differences were found between NZ and US children in corresponding age groups. While the mean for US 6-year-olds was significantly different from all other US age groups, the mean for NZ 6-year-olds was different only from the mean for NZ 10-year-olds.

3. Effects of Independent Variables within the VSM subtest

Age, gender and test order variables were examined on the VSM test, as well as on the RS and CO tests. Main effects were found for age and order on all three tests. Multiple comparisons of means for age on the VSM test and for order on all three tests are presented.

A significant effect was found for age, $F(4,177) = 18.61$, $p=.0001$. Means for Categories 1 (lsm=20.69) and 2 (lsm=22.12) were significantly different from Categories 3 (lsm=25.92), 4 (lsm=25.98), and

5 (lsm=29.11), p<.005. The mean for Category 3(lsm=25.92) was significantly different from Category 5 (lsm=29.11), p<.005 .

Although a main effect was found for order on the VSM test, F(5,177) = 2.54, p<.05, comparisons of means for each order were not significantly different at p<.0033. Main effects for order also were found on the RS and CO tests, F(5,178) = 2.29, p<.05 and F(5,178) = 3.20, p<.01, respectively. Post hoc comparisons of means showed that significant differences were found between order 2 (lsm=25.44) and Order 5 (lsm=21.86), p<.0033 on the RS test and between Order 3 (lsm=35.16) and order 6 (lsm=29.62), p<.0033 on the CO test. These results indicate that scores were significantly lower when the test was given first than when it was given last only for the RS and CO tests, but not for the VSM test.

Discussion

1. Comparison of NZ and US Performance

The performance of NZ and US children on the VSM was compared for corresponding age groups in the range of 6 - 10 years. Performance by NZ children aged between 7 and 10 years generally was not different to US children of the same age, however NZ 6-year-olds performed significantly better than US 6-year-olds. This finding is in contrast to the earlier studies by St George (1972) and Gronwall (1980). In those studies, the performance of NZ children was either the same or worse than US children of the same age.

While in the US sample performance of 6-year-olds was significantly different from every other age group, in the NZ sample performance of 6-year-olds was only significantly different from 10-year-olds. Performance by NZ children did not start to improve significantly until 9 to 10 years of age; scores by children in the younger age groups were in a similar range. In contrast, performance by US children started to level off in the older age groups.

A number of factors may account for these differences. The selection of the participants used in this study was based on different criteria than in the original standardisation sample. Participants in the original sample met specific criteria and were classified as "average" children. No exclusionary criteria were set for the present study. Participants were chosen on the basis of their willingness to participate and consent from parents and teachers. The range of educational abilities, language skills, and ethnic background was varied. It is possible that, given the small samples in each age group, a preponderance of certain characteristics may have skewed performance in a particular direction. Furthermore, the larger number of participants in the US study may account for the differences in performance by US and NZ children.

Changes in the school curriculum over the past 23 years since the original standardisation of the VSM, different exposure to school and life experiences (e.g., improved educational standards and facilities, television, computers), and improved intelligence scale ratings may have contributed to different results in this study. This also may account for the differences between this study and the earlier work by St George (1972) and Gronwall (1980). Furthermore, the norms may no longer be applicable for US children.

The improved performance by NZ six-year-olds however fits well with the fact that NZ children generally start school a year earlier than US children. It is possible that NZ children may be more used to manipulating or dealing with symbols (i.e., letters and numbers). However, it would be necessary to investigate preschool curriculum in the US before drawing conclusions on this point. It would have been useful to test NZ five-year-olds and to compare their performance with US five-year-olds.

2. Developmental Trends

Significant age effects were found on the VSM subtest and there was an increasing linear trend in performance on the test. Ceiling effects were not found on the VSM, however a number of participants came close to achieving the maximum score. Although Gronwall (1980), using a small sample, did not find ceiling effects for adults on the VSM, it is possible that the test may not be suitable for older participants.

Comparisons of means identified two or three main age groupings. There were three separate groupings on the VSM test: 6 - 9 years; 10 - 11 years; and 12 - 16 years. Hitch and Halliday (1983) found that children do not spontaneously rehearse visually presented items until they approach 10 years of age. Furthermore, Rose, Cundick, and Higbee (1983) found that inadequate task performance in reading comprehension was related, in part, to the ineffective use of memorisation strategies (verbal rehearsal and visual imagery) in learning. The increase in performance on the VSM by children aged 10 years and older therefore may be due to the development of more effective or more sophisticated strategies such as rehearsal.

Children of different ages also may have different expectations about the task and the consequences of their performance. Younger children treated the testing as a game and fun activity, whereas many of the older

children viewed the session as a test of their general abilities or intelligence. This may have encouraged older participants to try harder or to stay on task for longer periods than the younger participants.

3. General Conclusions

This study found that the US norms on the ITPA-Revised VSM subtest are, with caution, applicable to NZ children aged between 7 and 10 years. Significant differences were found between NZ and US children aged six years. The overall performance by NZ children was significantly better than US children. There are many factors that may account for the observed differences. For example, this study did not use the specific selection criteria as used in the original standardisation sample, but included children from different cultural backgrounds and presumably with a broader range of educational abilities. The difference between the US and NZ sample sizes, as well as different exposure to school and life experiences since the original standardisation of the VSM, also may contribute to variation in performance.

References

A.P.A., American Educational Research Association and National Council on Measurement in Education. (1974). *Standards for educational and Psychological tests.* Washington: A.P.A., Inc.

Amoriell, W.J. (1979). Reading achievement and the ability to manipulate visual and auditory stimuli. *Journal of Learning Disabilities, 12,* 64-66.

Ballard, K. (1988). Interpreting Stanford-Binet and WISC-R IQs in New Zealand: The need for more than caution. In M. Olsen (Ed.), *Mental testing in New Zealand: Critical and oppositional perspectives* (pp. 113-127). Dunedin: University of Otago Press.

Crothers, C. (1992*). A social profile of Auckland suburbs. 1991.* Auckland: Department of Sociology, University of Auckland (Working papers in comparative sociology, No. 22 - 1992).

Gronwall, D. (1980). Age and the visual sequential memory subtest of the ITPA. *New Zealand Psychologist,* 2, 76-78.

Hirshoren, A.A. (1969). A comparison of the predictive validity of the Revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities. *Exceptional Children, 35,* 517-521.

Hitch, G.J., & Halliday, M.S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society London B, 302,* 325-340.

Kass, C.E. (1966). Psycholinguistic disabilities of children with reading problems. *Exceptional Children, 8,* 533-539.

Kirk, S.A., & Kirk, W.D. (1971). *Psycholinguistic learning disabilities: Diagnosis and remediation.* Urbana: University of Illinois Press.

Kirk, S.A., & McCarthy, J.J. (1961). The Illinois Test of Psycholinguistic Abilities - An approach to differential diagnosis. *American Journal of Mental Deficiency, 66,* 399-412.

Kirk, S.A., McCarthy, J.J., & Kirk, W.D. (1968). *Examiner's Manual. Illinois Test of Psycholinguistic Abilities* Urbana: University of Illinois Press.

Mittler P., & Ward, J. (1970). The use of the Illinois Test of Psycholinguistic Abilities on British 4-year-old children: A normal and factorial study. *British Journal of Educational Psychology, 40,* 43-54.

Paraskevopoulos, J.N., & Kirk, S.A. (1969). *The development and psychometric characteristics of the Revised ITPA.* Urbana: University of Illinois Press.

Pavlidis, G.Th., & Fisher, D.F. (1986). *Dyslexia: Its neuropsychology and treatment.* Chichester: John Wiley & Sons.

Preston, J.M. (1984). ITPA-Revised. In D.J. Keyser & R.C.Sweetland (Eds.), *Test critiques* (Vol. 1, pp. 354-364). Missouri:Test Corporation of America.

Rose, M.C., Cundick, B.P., & Higbee, K.L. (1983). Verbal rehearsal and visual imagery: Mnemonic aids for learning disabled children. *Journal of Learning Disabilities, 16,* 352-354.

Salvia, J., & Ysseldyke, J.E. (1991). *Assessment.* Haughton Mifflin: Boston.

St.George, R. (1972). Maori and European psycholinguistic abilities: A resolution of results in conflict with similar studies. *Australian Journal of Psychology, 24,* 9-11.

Teasdale, G.R., & Wray, R.H. (1975). The use of the Revised ITPA with Australian children: A preliminary narrative study. *Australian Psychologist, 10,* 13-20.

Wilson, B.C. (1992). An approach to the neuropsychological assessment of the preschool child: An hypothesis-testing model. (abstract of conference workshop). *Journal of clinical and Experimental Neuropsychology, 14(1),* 4.

Zaidel, E. (1979). Performance on the ITPA following cerebral commisurotomy and hemispherectomy. *Neuropsychologia, 17,* 259-280.

Address for Correspondence

Sandra A Heriot

Department of Psychology

University of Waikato

Private Bag 3105, Hamilton

New Zealand

Email: sah10@waikato.ac.nz