

**EARLY LITERACY
RESEARCH PROJECT**

Time 2 Assessments of Children's Literacy Skills

A Report of Findings for the
Ministry of Education

James W. Chapman, Alison W. Arrow, and
William E. Tunmer
Massey University
October, 2015



EARLY LITERACY RESEARCH PROJECT

**Report for the Ministry of Education of Findings from the Time 2 (June/July 2015)
Assessments of Children's Literacy Skills**

James W. Chapman, Alison, W. Arrow and William E. Tunmer

Massey University

Palmerston North

October, 2015

Summary

This report presents data on the demographic information about the students and data for Time 2 assessments conducted during June and July 2015.

Despite some degree of attrition (17%), including one comparison school that decided to withdraw from the project, children in the intervention and comparison groups are generally similar in terms of age, gender and decile rank of their school. Moreover, as expected there were no statistically significant differences between the two groups on any of the Time 2 assessment measures.

An examination of the distribution characteristics of scores for many measures revealed skewed distributions, but these were not as marked as observed for the Time 1 assessments. As anticipated, differences as a function of school decile band remained in these Time 2 data. On all variables, children in the higher decile schools (deciles 8-10) made greater progress than children in the lower decile schools (deciles 1-3). Compared to the Time 1 data, there appears to be a widening of the gap in performance between low and high decile children on a number of variables. If the project is successful in one of its major goals, children in intervention low decile schools should outperform their counterparts in the comparison schools, and reach or at least approximate the literacy development of children in higher decile schools.

Introduction

The main goal of this early literacy research project is to improve the literacy learning outcomes of New Entrant children during and beyond their first year of schooling. A particular focus is on children from lower decile schools, and children from Maori and Pasifika backgrounds.

Our approach in this research project is to work with teachers of New Entrant children in schools randomly selected from the southern half of the North Island (see Chapman, Arrow & Tunmer, 2015 for a description of the sampling procedure). Our work with teachers is focussed on supplementary teaching strategies that are designed to enhance regular classroom literacy instruction, with a particular emphasis on the development of language-related and word-level decoding skills that are essential for successful literacy acquisition.

In our report on the Time 1 assessment data (Chapman et al., 2015) we provided data to show that children in the intervention and comparison groups are generally similar in terms of age, gender and decile rank of their school. There were no statistically significant differences between the two groups on any of the Time 1 assessment measures. We concluded that based on these data, the children in both groups are essentially equivalent in terms of these important baseline literacy variables.

In this report, we present the results of children's assessments at Time 2, which was during the period of June and July 2015. By this time, teachers from the intervention schools had participated in two professional learning and development workshops over a total of three days.

We do not expect differences between the intervention and comparison groups to be observable at such a relatively early stage in the project. Changes in teacher knowledge and teaching practice take time. Some research indicates that changes take some time to become embedded in daily instructional approaches when teachers are engaged in learning activities that present strategies that are different from their current teaching approaches.

Nonetheless, we thought it useful to undertake a repeat of the Time 1 assessments (February/March) at the middle of the year and before the end of Term 2 (June/July) to examine whether both intervention and comparison groups continued with similar literacy performance levels. We also examined literacy development in relation to school decile ranking to determine whether initial differences between children from low and high decile schools remained constant during the first 5 months of their schooling.

Sample

At the start of the project in February 2015, 38 schools confirmed their willingness to participate in the research. Since then, one of the comparison schools made a decision to

withdraw from the project prior to the administration of the Time 2 assessments. This withdrawal resulted in a reduction of 22 comparison children from 158 to a potential of 136.

The number of children assessed at Time 2 declined from 359 at the start of the project to 297 (17.3% decrease). The declines were due mainly to absenteeism at the time the assessments were administered or departure from the school. The intervention group reduced by 30 children, from 202 to 172 (14.9% decrease). The comparison group decreased by 34 children, from 159 to 125 (a 21.2% decrease, inflated by the 22 children at the comparison school who were withdrawn from the project).

Decile rankings of schools were grouped into three bands: low = deciles 1 to 3; medium = deciles 4-7; high = deciles 8-10. At the Time 2 assessment point, the spread across these three decile bands remained more even for the intervention schools than the comparison schools. These data are presented in Table 1.

Table 1. Distribution of project participants by group and school decile band.

Decile Band	Group	
	Intervention	Comparison
	Percent (n)	Percent (n)
1-3	35.5 (61)	67.2 (84)
4-7	28.5 (49)	6.4 (8)
8-10	36.0 (62)	26.4 (33)

Sample Characteristics

Of the 38 schools that agreed to participate in the project, 24 had been randomly assigned to the intervention group, and 14 to the comparison group. A total of 62 teachers of New Entrant children were identified: 38 in the intervention schools and 24 in the comparison schools. These numbers fluctuate as teachers come and go for various personal or professional reasons. At the first of the scheduled intervention group teacher professional development workshops, 45 teachers attended from the 24 intervention schools. At the second workshop in May, 46 teachers attended from the intervention schools.

Time 2 assessment data were collected during June and July 2015 from 305 New Entrant/Year 1 children. Of these, 172 (56%) were in intervention schools, and 133 (44%) were in comparison schools.

Age data reveal that the mean age of the sample at end of Term 2 (3 July, 2015) was 65.56 months ($SD = 1.81$), which is 5 years, 6 months; the median age was 65 months, and the mode was 65 months. Clearly, the large majority of children were around 5 years 5 months of age. There is no statistically significant difference between the mean ages for the Intervention and Comparison groups: 65.6 ($SD = 1.68$) and 65.6 ($SD = 1.97$) respectively.

In terms of gender, 54% of the children were boys and 46% were girls. However, there was a marginal imbalance for children in the intervention group: 56% (96) were boys compared to 44% (76) girls. The gender breakdown in the comparison group was generally more even: 53% (70) boys and 47% (63) girls.

Regarding ethnic background of the participating students, we anticipate providing these data in the report on Time 3 assessment data when we expect to have obtained these and recorded them in the database.

Assessments

The following assessments, first undertaken during February and March, 2015, were repeated during June and July 2015.

Letter Identification

Research indicates that letter name knowledge and letter sound knowledge are important aspects of initial literacy acquisition. We have kept these two aspects of letter identification separate, because they differentially predict pathways to new word learning (Arrow, 2012).

Letter name and letter sound knowledge were assessed for both upper case and lower case letters, using the Letter Identification task in the Diagnostic Survey (Clay, 1985). Children were asked to name each letter and to say the sound the letter represented for 26 upper case and 28 lowercase letters, two of which appeared in varying fonts. Scoring was based on the number of letters correctly identified by name and by sound.

Word Recognition

Word recognition refers to the fluent, rapid reading of words as they appear. Such words are usually known as *sight words*. These words include the high-frequency words that children learn to read first, as well as any word that is read with automaticity. We used one of the *Ready to Read* test lists (Clay, 2002). These tests comprise 45 words of the most frequently occurring words in the 12 "little" books of the *Ready to Read* series. We administered the first 15 words in one of the lists. Scoring was based on the number of words read correctly by each child.

Invented Spelling

Invented spelling was assessed by having children write 18 words that were read aloud by the research assistant. The 26 (lower case) letters of the alphabet were displayed across the top of the children's response sheets. Each word that children wrote down received a score from 0 to 4. Maximum points were awarded if the sounds in the word were represented with letters, although unconventionally (e.g., *kik* for *kick*, *fil* for *fill*, *sid* for *side*). Two points were awarded if more than one phoneme (but not all) was represented with phonetically related or conventional letters (e.g., *sd* for *side*, *lup* for *lump*). One point was awarded where the initial phoneme was represented with the correct letter (e.g., *f* for *fat*). Children were also asked to identify the sounds in the words that were read aloud. The total number of possible points for letters and sounds was 72 each.

Phonological Processing

Phonological processing was assessed using the Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2: Wagner, Torgesen, Rashotte, & Pearson, 2013). This test is normed in the United States for use with people from 4 years to 25 years. The CTOPP-2 is used to help evaluate phonological processing abilities as a prerequisite to reading fluency. We administered three of the subtests: elision, blending and matching. Elision measures the ability to remove phonological segments from spoken words to form other words. There are 34 items in this test, with discontinuation occurring when each child missed three consecutive items. Blending Words measures the ability to synthesize sounds to form words. There were 33 items in this section; again, discontinuation occurred following three consecutive missed items. Sound Matching measures the ability to select words with the same initial and final sounds. This section comprised 26 items; testing was discontinued following three missed items.

Results

Score Distributions

We examined the distribution of scores on the Time 2 variables for the total sample. Whereas at Time 1 nine of the 12 variables showing modal scores of 0, on this assessment occasion, six of the 11 variables (the British Picture Vocabulary Test was not administered on this occasion) showed modal scores of 0: Letter ID lower case sound, CTOPP elision, Clay word test and Clay phonemes, and the two measures of invented spelling. Reflecting the uneven development in these variables upper and lower case sounds showed bi-modal distributions; upper and lower case names showed significant negatively skewed distributions (many children approached the ceiling on these tasks); the Clay tests and invented spelling showed positively skewed distributions. However, CTOPP blending and matching, along with invented spelling sounds, showed distributions that were close to normal. Because of questions regarding the veracity of the marking of the phonemes test based on the Clay words, we did not continue with analyses of this variable.

Group Comparisons

We performed simple *t*-tests to test the hypothesis that there would be no significant differences in mean scores between the intervention and comparison groups on the Time 2 assessment variables. The hypothesis was supported for each of the variables. Means and standard deviations are presented in Table 2.

Table 2. Means and standard deviations for intervention and comparison students at Time 2.

Assessments	Intervention Group			Comparison Group		
	Mean	SD	n	Mean	SD	n
Letter ID UC Name	20.50	7.20	172	21.26	7.21	133
Letter ID LC Name	20.73	7.80	172	22.00	7.91	133
Letter ID UC Sound	16.99	8.73	172	18.77	8.84	133
Letter ID LC Sound	17.97	9.12	172	19.71	9.53	133
Clay Word Test	5.29	4.39	172	6.09	4.25	133
Invented Spelling	2.15	2.70	172	2.59	3.03	133
Invented Spelling Sounds	27.30	20.55	172	28.70	20.45	133
CTOPP Elision	9.13	6.74	172	9.50	6.19	133
CTOPP Blending	12.01	6.44	172	11.87	6.47	133
CTOPP Matching	12.20	6.69	172	13.36	6.77	133

Decile Band Comparisons

We examined scores for the 10 variables in terms of decile bands for the total sample. These comparisons were conducted by means of one-way analyses of variance (ANOVAs). All variables showed statistically significant effects.

The means and standard deviations for the 10 variables are presented in Table 3. These data show that mean scores for children in the low decile band were lower than for children in the high decile band for all 10 variables. The low decile group was significantly lower than the middle decile group on only one variable, namely, CTOPP elision test.

However, the middle decile group attained lower scores than the high decile group on three variables: letter sound upper and lower and invented spelling sounds.

Table 3. Means and standard deviations for Time 2 variables as a function of school decile band.

Variables	Low Decile Band		Middle Decile Band		High Decile Band	
	Mean	SD	Mean	SD	Mean	SD
Letter UC Name*	19.81 ^a	7.78	20.72	7.63	22.21	5.96
Letter LC Name*	20.16 ^a	8.54	21.35	7.81	22.71	6.82
Letter UC Sound**	16.81 ^{ac}	8.97	15.72	9.92	20.23	7.28
Letter LC Sound**	17.75 ^{ac}	9.53	16.81	10.20	21.14	7.87
CTOPP elision**	7.40 ^{ab}	6.18	10.42	7.20	11.39	5.95
CTOPP blending**	10.51 ^a	6.18	11.84	6.40	14.11	6.48
CTOPP matching**	11.37 ^a	6.48	12.51	6.85	14.66	6.79
Clay word**	4.91 ^a	4.18	5.34	4.13	6.64	5.54
Invented spelling**	1.60 ^a	2.60	2.14	2.90	3.45	2.83
Invented spelling sounds**	22.63 ^{ac}	18.84	22.45	2.97	35.62	19.14

* Statistically significant 1-way ANOVA, $p < .05$

** Statistically significant 1-way ANOVA, $p < .01$

^a Low decile group significantly lower than high decile group

^b Low decile group significantly lower than middle decile group

^c Middle decile group significantly lower than high decile group

UC = upper case; LC = lower case

CTOPP = Comprehensive Test of Phonological Processing

We also examined the distributions of scores in terms of school decile band for all variables. This examination was undertaken because of the large numbers of students who scored 0 on many of the measures. We calculated the percentages of students scoring 0 or 1 on each of the 10 variables. These percentages are presented in Table 4.

Table 4. Percentages of students obtaining scores of 0 or 1 on each of the listed variables, as a function of school decile band.

Variables	Low	Middle	High
Letter UC Name	4.1	1.8	0.0
Letter LC Name	6.2	1.8	0.0
Letter UC Sound	11.7	21.0	3.2
Letter LC Sound	14.5	17.5	4.3
CTOPP elision	26.4	10.5	8.4
CTOPP blending	9.0	5.3	3.2
CTOPP matching	7.6	3.6	4.2
Clay word	29.9	23.7	21.5
Invented spelling	69.6	61.4	36.8
Invented spelling sounds	17.3	28.1	7.4

The children in the high decile band schools have generally progressed more rapidly than those in the low or middle decile bands. These results are, of course, consistent with the data presented in Table 3 from the analyses of variance. There is more variation in low scores between the low and middle decile children, with fewer low decile than middle decile children showing scores of 0 or 1 on some variables (viz., letter sounds—upper and lower case, Invented spelling sounds).

To provide a visual comparison of results for Time 1 and Time 2 assessments in relation to decile band, we present a series of line graphs in the Appendix. These graphs illustrate the “gap” in initial performance between high and low decile children. The gap looks to close slightly for Letter Name (upper and lower case); the gap appears to be consistent (parallel) for Letter Sound (upper and lower case) and CTOPP Matching; and, there appears to be a widening between high and low decile children for CTOPP Elision and Blending, Clay Word, and the two measures of Invented Spelling (words and sounds).

Discussion

Overall, findings from the Time 2 assessments indicate children in the intervention and comparison groups are generally similar in terms of age, gender and decile rank of their school, despite the attrition from the project that has taken place to date. Further, no

statistically significant differences between the intervention and comparison students have emerged at this stage of the project. This finding is consistent with our expectations.

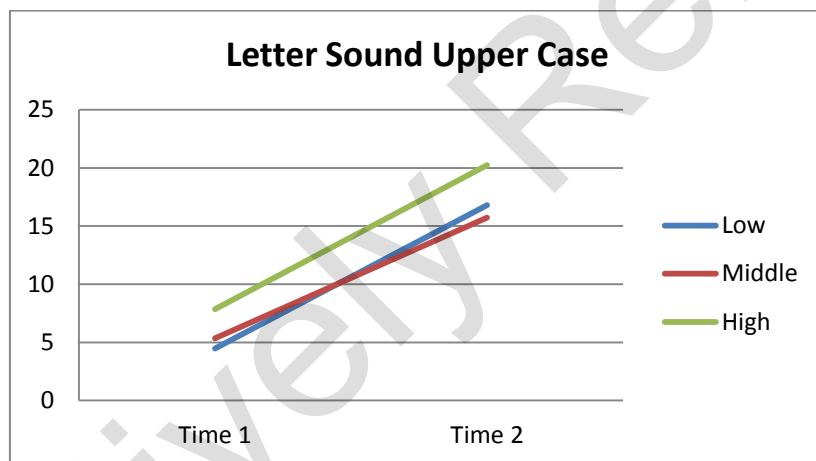
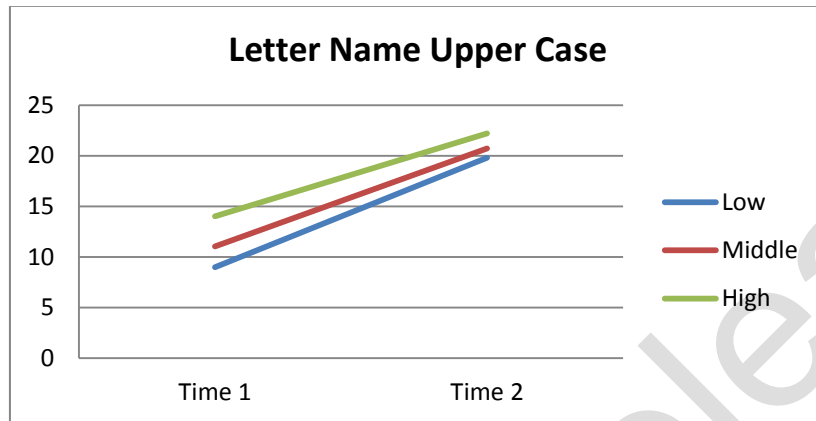
Finally, as expected, significant differences as a function of school decile band are apparent in these Time 2 data. The low decile group performed at lower levels than the high decile group on all variables, and lower than the middle decile group on one variable. We anticipate that children from intervention low decile schools should outperform their counterparts in the comparison schools, and reach or at least approximate the development of children in higher decile schools.

References

- Arrow, A.W. (2012, July). *Alphabet knowledge in the acquisition of early orthographic representations*. Paper presented at the annual meeting of the Society for the Scientific Studies in Reading, Montreal, Canada.
- Chapman, J.W., Arrow, A.W., & Tunmer, W.E. (2015). *Early literacy research project: Report to the Ministry of Education of Findings from the Time 1 (February/March 2015) Assessments of Children's Literacy Skills*. Palmerston North: Massey University.
- Clay, M.M. (1985). *The early detection of reading difficulties*. Auckland: Heinemann.
- Clay, M.M. (2002). *An observation survey of early literacy achievement* (2nd. Ed.). Auckland, Heinemann.
- Wagner, R.K., Torgesen, J.K., Rashotte, C.A., & Pearson, N.A. (2013). *Comprehensive test of phonological processes* (2nd. Ed.). Austin, TX: Pro-Ed.

Appendix

Line Graphs Comparing Time 1 with Time 2 Mean Scores as a Function of School Decile Band¹



¹ Low decile band = deciles 1-3; middle decile band = deciles 4-7; high decile band = deciles 8-10

