Research Summary on InitiaLit-Foundation

We have carried out numerous informal field trials during the development of InitiaLit – Foundation, most of which have not involved formal data collection. They were carried out to see how well the programs worked in the real world of classrooms and changes were made on the basis of teacher feedback and our observations.

We have, however, carried out three preliminary data based trials of the draft program so far. Two trials carried out in 2016 compared the early literacy skills of children who received instruction in the program to a comparison group of students who received different instruction. One trial involved five Foundation (Kindergarten) classes in Sydney and the other trial in Perth involved two Foundation classes. Both schools had students from average socioeconomic backgrounds and had a high proportion of students with a language background other than English. In the Sydney trial, two classes (35 students) were allocated to the treatment condition (i.e., those receiving InitiaLit) and three classes (37 students) served as a comparison group. In Perth, there was one class each in treatment (InitiaLit) and comparison conditions (27 and 17 students) respectively. Note that we are not claiming anything more than preliminary findings from these trials. They were not randomised control trials. The initial, preliminary, results for these efficacy studies were presented in a conference paper (Wheldall et al., 2017).

In both studies in 2016, students were assessed on the Foundations of Early Literacy Assessment (FELA; Neilson, 2016) and the York Assessment of Reading for Comprehension (YARC) Early Reading (Hulme et al., 2012). The FELA, devised by Dr Roslyn Neilson, assesses the phonemic awareness (PA) skills required for literacy development of children in the first few years of school. It identifies children who have difficulties with PA and indicates their areas of strength and weakness. The YARC Early Reading assesses early reading skills including alphabetic knowledge, single word reading and phoneme awareness.

In the Sydney school, the comparison was with 'business as usual' instruction for these classes which in this instance was essentially a Whole Language approach. A systematic program of instruction in phonics did not form part of the curriculum. In the Perth school, InitiaLit was taught by a second year out teacher, teaching his own class for the first time, and the comparison was with a class receiving exemplary, explicit phonics instruction using the Letters and Sounds program by a very experienced teacher who had been trained in explicit teaching and direct instruction.

In both schools, students in both conditions were assessed prior to (pre-test) and following 16 weeks of InitiaLit instruction (mid-test). In the Sydney school, students were assessed again after a further 12 weeks of instruction (28 weeks in total; post-test). In the Perth school, students were assessed again after a further 18 weeks of instruction (34 weeks in total; post-test). All tests were administered on all three occasions except that the FELA was not given to the students in Sydney at post-test.

In Sydney, our results showed that for almost all measures at mid-test (after 16 weeks) and post-test (28 weeks) the experimental group performed significantly better than the comparison group (with moderate or large effect sizes [partial eta squared] – see Tables 1 and 2). These analyses were carried out using raw score data. These gains appear to be cumulative, as we would expect. For Letter Sound Knowledge, the gap increased between the treatment and comparison classes in Sydney. Figure 1 provides an illustration of the treatment group compared to the comparison group for Letter Sound Knowledge expressed as standard scores (adjusted for any pre-test differences) which take increasing age over the intervention into account. (Note that this is a conservative estimate of growth for the InitiaLit group because seven students scored above the top standard score norms provided by the YARC [130], whereas no students in the comparison group scored above the norms.) The graph clearly shows how the treatment group continued to show good growth in letter sound knowledge skills in the second half of the program.

Table 1. Sydney school means, standard deviations (raw scores) and comparison of groups at mid-test, controlling for variation at pre-test.

Measure	Group	N	Pretest		Mid Test		Sig	Partial Eta Sq
			M	SD	M	SD		
Letter sound knowledge	Comparison	37	4.86	6.59	16.35	7.71	0.012	0.089
	InitiaLit	35	7.60	8.15	21.11	5.06		
Early Word Recognition	Comparison	37	0.14	0.42	5.95	5.39	0.263	0.018
	InitiaLit	35	2.77	6.38	10.23	8.53		
Phoneme Awareness	Comparison	37	2.57	3.35	8.05	5.32	0.011	0.090
	InitiaLit	35	3.94	3.61	11.66	4.28		
FELA	Comparison	37	25.73	19.91	67.27	26.51	0.000	0.248
	InitiaLit	35	37.71	24.02	99.26	23.25		

Note: Partial Eta Squared is an effect size calculation (% of variance explained). A small effect is.01 or 1%; a medium effect is .06 or 6%; and a large effect is .138 or 13.8%.

Table 2. Sydney school means, standard deviations (raw scores) and comparison of groups at post-test, controlling for variation at pre-test.

Measure	Group	N	Pretest		Post Test		Sig	Partial Eta Sq
			M	SD	M	SD		
Letter sound knowledge	Comparison	37	4.86	6.59	21.43	6.30	0.000	0.372
	InitiaLit	35	7.60	8.15	29.29	2.92		
Early Word Recognition	Comparison	37	0.14	0.42	11.89	8.76	0.011	0.091
	InitiaLit	35	2.77	6.38	18.54	7.14		
Phoneme Awareness	Comparison	37	2.57	3.35	11.35	5.19	0.011	0.091
	InitiaLit	35	3.94	3.61	14.89	4.18		

Note: Partial Eta Squared is an effect size calculation (% of variance explained). A small effect is.01 or 1%; a medium effect is .06 or 6%; and a large effect is .138 or 13.8%.
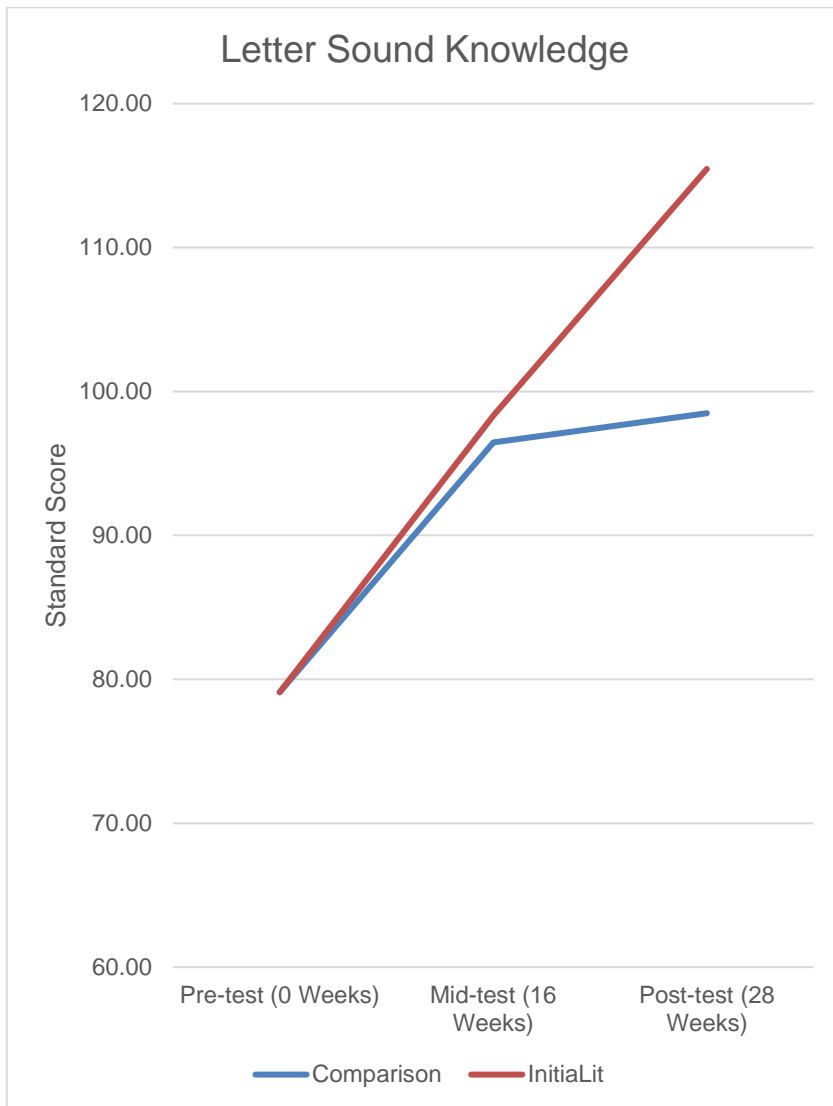
Figure 1. Adjusted standard scores for the two groups (Sydney)

At pre-test 78% and 74% in comparison and InitiaLit groups respectively scored in the bottom quartile (bottom 25%) for letter sound knowledge. No students scored in the top quartile in the comparison group and only (3%) did so in the InitiaLit group. At the final post-test, however, whereas 22% of students in the comparison group still scored in the bottom quartile, only 3% did so in the InitiaLit group. Moreover, whereas 71% scored in the top quartile in the InitiaLit group, only 11% did so in the comparison group. InitiaLit had almost eliminated the tail of low-progress readers and had pushed a far higher proportion into the top quartile.

Turning now to the Perth school, our preliminary results showed that the InitiaLit group performed as well as, but not significantly differently from, the comparison group (see tables 3 and 4). In fact, at mid-test, after 16 weeks, the comparison group were performing better than the InitiaLit treatment group for Letter Sound Knowledge. After a further 18 weeks, however, this was reversed with the InitiaLit group now ahead. This is made clear in Figure 2 showing progress over the three test occasions for the adjusted standard scores for Letter Sound Knowledge. (Note that the growth estimated by standard scores is also conservative for the InitiaLit group in this school because eight students scored above the highest standard score provided in the YARC normative data, compared to two in the comparison group).

Table 3. Perth school means, standard deviations (raw scores) and comparison of groups at mid-test, controlling for variation at pre-test.

Measure	Group	N	Pretest		Mid Test		Sig	Partial Eta Sq
			M	SD	M	SD		
Letter sound knowledge	Comparison	17	10.94	7.77	25.12	2.29	0.000	0.351
	InitiaLit	27	10.22	8.94	21.30	3.50		
Early Word Recognition	Comparison	17	0.24	0.66	5.71	3.74	0.228	0.035
	InitiaLit	27	0.52	1.42	7.85	5.26		
Phoneme Awareness	Comparison	17	6.06	2.56	11.24	2.80	0.234	0.034
	InitiaLit	27	7.00	5.39	12.70	3.74		
FELA	Comparison	17	36.53	12.41	98.06	11.74	0.794	0.002
	InitiaLit	27	42.59	27.00	101.85	17.88		

Note: Partial Eta Squared is an effect size calculation (% of variance explained). A small effect is .01 or 1%; a medium effect is .06 or 6%; and a large effect is .138 or 13.8%.

Table 4. Perth school means, standard deviations (raw scores) and comparison of groups at post-test, controlling for variation at pre-test.

Measure	Group	N	Pretest		Post Test		Sig	Partial Eta Sq
			M	SD	M	SD		
Letter sound knowledge	Comparison	17	10.94	7.77	30.06	1.71	0.031	0.109
	InitiaLit	27	10.22	8.94	30.93	1.33		
Early Word Recognition	Comparison	17	0.24	0.66	17.47	5.10	0.700	0.004
	InitiaLit	27	0.52	1.42	18.48	4.71		
Phoneme Awareness	Comparison	17	6.06	2.56	15.12	3.46	0.139	0.053
	InitiaLit	27	7.00	5.39	16.74	2.90		
FELA	Comparison	17	36.53	12.41	123.59	8.05	0.100	0.064
	InitiaLit	27	42.59	27.00	128.15	7.46		

Note: Partial Eta Squared is an effect size calculation (% of variance explained). A small effect is .01 or 1%; a medium effect is .06 or 6%; and a large effect is .138 or 13.8%.
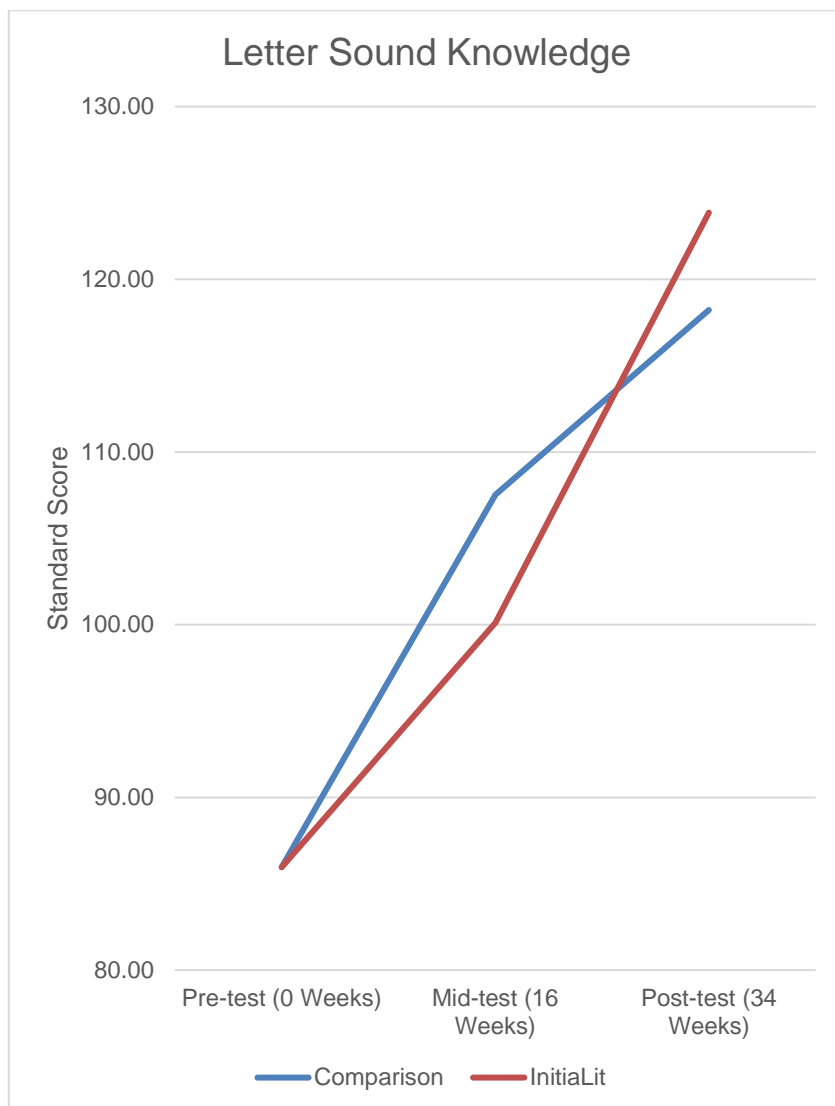
Figure 2. Adjusted standard scores for the two groups (Perth)

For the comparison and InitiaLit group at pre-test, 59% and 67% respectively scored in the bottom quartile for letter sound knowledge; 12% and 4% respectively scored in the top quartile. At final post-test, no students in either group scored in the bottom quartile but 85% of students in the InitiaLit group scored in the top quartile compared to 71% in the comparison group. Given that a smaller proportion of students scored in the bottom quartile and a higher proportion scored in the top quartile in the comparison group at pre-test, the results for the InitiaLit group suggests that they were making greater progress overall.

In 2017, another trial was conducted in primary schools in Sydney with all students in Foundation. Of the students receiving instruction in this program, 63 students were assessed across the school year on measures of early reading skills. One of the school's students came from average socioeconomic backgrounds and no socioeconomic status information was available for the other school. Both schools had a high proportion of students with a language background other than English.

Students were assessed prior to the commencement of instruction (pre-test), following approximately 20 weeks of instruction (mid-test) and again following another 17 weeks of instruction (post-test). The measures used in this trial differed slightly from those used in the previous year's trials. On all three testing occasions, students were assessed on the York Assessment of Reading for Comprehension (YARC) Early Reading (Hulme et al., 2012). At mid-test and post-test, students were also assessed on an experimental measure of decoding fluency, the Wheldall Assessment of Reading Nonwords (WARN), and at post-test only, students' phonological recoding skills were assessed using the Martin and Pratt Nonword Reading Test (Martin & Pratt, 2001). Over the course of the year, the schools involved both completed all 126 lessons of the program.

Students made significant gains over the year with large effect sizes on all measures of early reading skills, including letter sound knowledge, word recognition and phoneme awareness, as shown in Table 5. Because these students were initially assessed at the beginning of their Foundation year when they would not be expected to have had any formal instruction in skills such as letter sound knowledge or word recognition, it is not surprising that these students made significant gains in these skills subsequent to reading instruction in the first half of Foundation. We would expect most children to make some gains regardless of the type of instruction they received. However, these gains were not only significant when measured over the first half of the year (see Table 6), but also over the second half of the year (see Table 7). There was also a significant gain with a large effect size made in decoding fluency over the second half of the year. This indicates that the program had a sustained impact on the children's acquisition of these skills and that the impressive gains are less likely to be solely due to their attendance to some form of instruction.

Table 5. Means (and standard deviations) and the resultant gains on measures of early reading skills (raw scores) for Foundation students over the entire year.

Literacy Variable	N	Raw Score Pre-test (sd)	Raw Score Post-test (sd)	Gain (sd)	t	p	Partial Eta Sq
Letter Sound Knowledge	63	7.33 (8.28)	30.35 (3.98)	23.02 (8.28)	22.05	<0.0005	0.887
Early Word Recognition	63	2.25 (5.71)	19.71 (7.77)	17.46 (7.33)	18.89	<0.0005	0.852
Phoneme Awareness	63	3.95 (4.68)	15.87 (4.24)	11.92 (4.50)	21.03	<0.0005	0.877
Decoding Fluency	62	Not assessed	10.43 (5.72)	N/A	N/A	N/A	N/A
Phonological Recoding	63	Not assessed	18.08 (8.30)	N/A	N/A	N/A	N/A

Note: Partial Eta Squared is an effect size calculation (% of variance explained). A small effect is.01 or 1%; a medium effect is .06 or 6%; and a large effect is .138 or 13.8%.

Table 6. Means (and standard deviations) and the resultant gains on measures of early reading skills (raw scores) for Foundation students during the first half of the year.

Literacy Variable	N	Raw Score Pre-test (sd)	Raw Score Mid-test (sd)	Gain (sd)	t	p	Partial Eta Sq
Letter Sound Knowledge	63	7.33 (8.28)	24.33 (4.65)	17.00 (6.70)	20.14	<0.0005	0.867
Early Word Recognition	63	2.25 (5.71)	13.13 (7.36)	10.87 (5.67)	15.23	<0.0005	0.789
Phoneme Awareness	63	3.95 (4.68)	12.43 (4.90)	8.48 (3.96)	16.99	<0.0005	0.823
Decoding Fluency	62	Not assessed	6.63 (4.43)	N/A	N/A	N/A	N/A

Note: Partial Eta Squared is an effect size calculation (% of variance explained). A small effect is.01 or 1%; a medium effect is .06 or 6%; and a large effect is .138 or 13.8%.

Table 7. Means (and standard deviations) and the resultant gains on measures of early reading skills (raw scores) for Foundation students during the second half of the year.

Literacy Variable	N	Raw Score Mid-test (sd)	Raw Score Post-test (sd)	Gain (sd)	t	p	Partial Eta Sq
Letter Sound Knowledge	63	24.33 (4.65)	30.35 (3.98)	6.02 (3.40)	14.06	<0.0005	0.761
Early Word Recognition	63	13.13 (7.36)	19.71 (7.77)	6.59 (4.60)	11.37	<0.0005	0.676
Phoneme Awareness	63	12.43 (4.90)	15.87 (4.24)	3.44 (3.17)	8.63	<0.0005	0.541
Decoding Fluency	62	6.63 (4.43)	10.60 (5.79)	3.97 (3.68)	8.48	<0.0005	0.54
Phonological Recoding	63	Not assessed	18.08 (8.30)	N/A	N/A	N/A	N/A

Note: Partial Eta Squared is an effect size calculation (% of variance explained). A small effect is .01 or 1%; a medium effect is .06 or 6%; and a large effect is .138 or 13.8%.

Although as part of this trial there is no comparison class which received the usual instruction provided by the teachers with which to compare these results, we can gain some meaningful indication of progress by comparing these results to the typical progress of students this age. We therefore compared students' results at pre-, mid- and post-test to other students of their age. Average (or mean) standard score results, which compare student performance with their same aged peers, were analysed and are presented in Table 8. If we consider the average range of performance to be between standard scores of 85 and 115 with the average score being 100, the students scored below average on measures of letter sound knowledge and phoneme awareness at pre-test. However, by the middle of the year they achieved a mean standard score that was in the average range for phoneme awareness and above the average score for letter sound knowledge. For the measure of word recognition, while the students' mean score was within the average range at pre-test, by mid-test, they scored above the average score for same aged peers. Furthermore, these estimates of average performance are conservative because at pre-test, up to 43% of the students scored below the range of standard scores provided by the test (and hence the mean score is overestimated) while at mid-test, only up to 7% of students scored below this range. Table 6 and Figure 3 below show the students' average progress in terms of standard scores on these three measures.

Table 8. Means (and standard deviations) on measures of early reading skills (standard scores) for Foundation students at pre-, mid- and post-test.

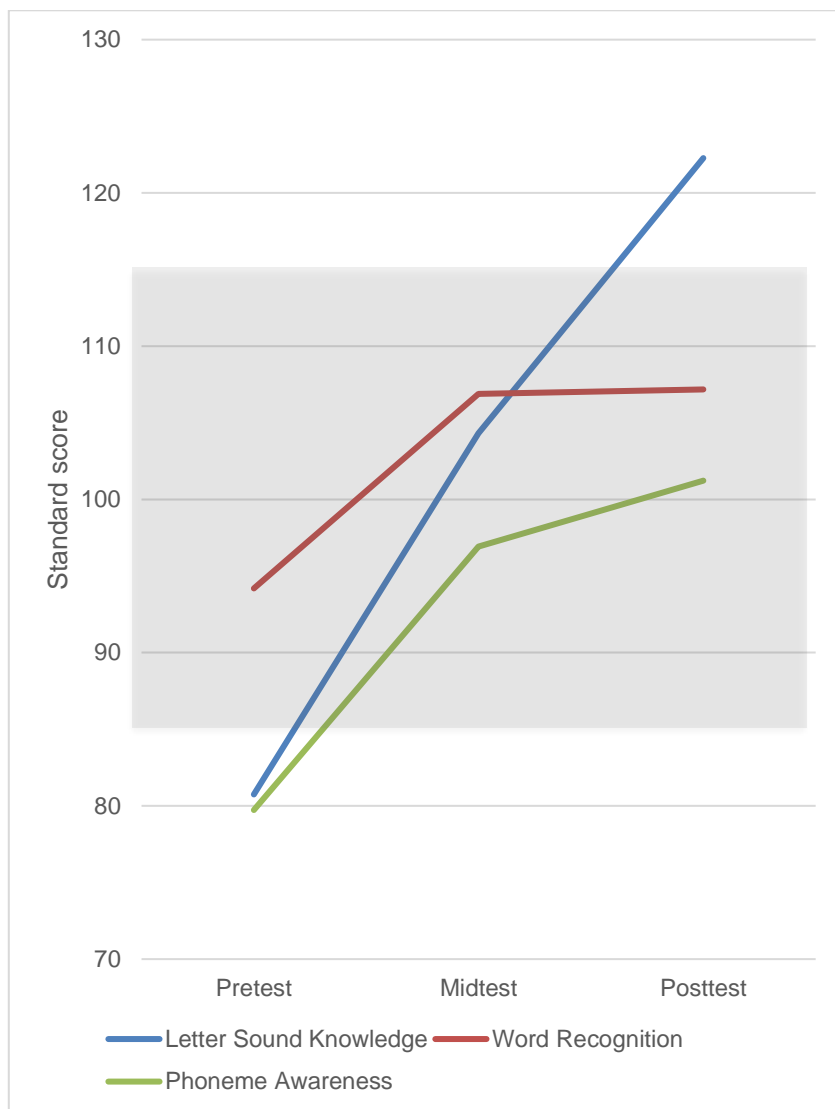| Literacy Variable	N	Standard Score Pre-test (sd)	Standard Score Mid-test (sd)	Standard Score Post-test (sd)
YARC Letter Sound Knowledge	63	80.75 (15.05)	104.32 (10.96)	122.27 (13.19)
YARC Early Word Recognition	63	94.19 (10.77)	106.89 (12.14)	107.17 (13.62)
YARC Phoneme Awareness	63	79.73 (12.81)	96.92 (12.36)	101.22 (12.70)
Phonological Recoding	34	Not assessed	Not assessed	113.12 (12.02)

Figure 3. Foundation students' performance relative to same aged peers on measures of early reading skills at pre-, mid- and post-test. The average range of student performance, 85 to 115 in standard scores, is indicated by the grey shading.

Further analysis revealed that there was a considerable shift of students out of the bottom quartile (bottom 25% of same aged students) to the average range (middle 50% of same aged students) and top quartile (top 25% of same aged students) between pre-, mid- and post-test as shown in Table 9 (also see Figures 4 to 7). At pre-test, 75% of students scored in the bottom quartile for letter sound knowledge and only 6% scored in the top quartile. At mid-test, only 10% of students remained in the bottom quartile and 41% of students scored in the top quartile. By post-test, while 6% remained in the bottom quartile, 89% of students scored in the top quartile. Similarly, on the measure of phoneme awareness, 79% of students scored in the bottom quartile and only 3% scored in the top quartile at pre-test. At mid-test, only 30% remained in the bottom quartile and 54% scored in the top quartile. By post-test, only 14% of students remained in the bottom quartile and 22% scored in the top quartile. This indicates that the majority of students had moved to the average range in phoneme awareness after approximately 37 weeks of instruction. For word recognition, 19% of students were in the bottom quartile and 8% were in the top quartile at pre-test. At mid-test, only 11% remained in the bottom quartile and 48% were now in the top quartile. By post-test, 16% scored in the bottom quartile and 51% were now in the top quartile. The measure of phonological recoding, which was only appropriate for use with these students at post-test, revealed a similar pattern of student results compared to others of the same age. At post-test, only 6% of students scored in the bottom quartile and the majority (65%) scored in the top quartile for this skill. These results indicate that InitiaLit-F may have helped to reduce the number of students who might have struggled to learn to read (those in the bottom quartile) while not limiting the

growth of higher performing students, as indicated by those moving from the average range to the top quartile.

Table 9. Foundation students performing in the bottom quartile (bottom 25% of students), average range and top quartile at pre-, mid- and post-test.

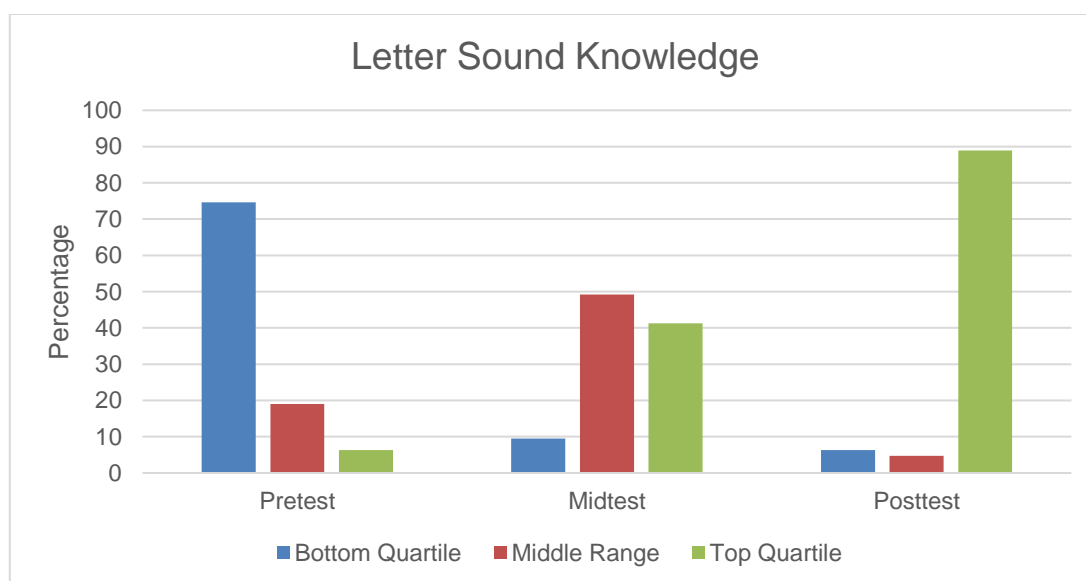|		Bottom Quartile	Average	Top Quartile
Letter Sound Knowledge	Pre-test	74.6%	19.0%	6.3%
	Mid-test	9.5%	49.2%	41.3%
	Post-test	6.3%	4.8%	88.9%
Word Recognition	Pre-test	19.0%	73.0%	7.9%
	Mid-test	11.1%	41.3%	47.6%
	Post-test	15.9%	33.3%	50.8%
Phoneme Awareness	Pre-test	79.4%	17.5%	3.2%
	Mid-test	30.2%	15.9%	54.0%
	Post-test	14.3%	63.5%	22.2%
Phonological Recoding	Post-test	5.9%	29.4%	64.7%

Figure 4. Percentages of students scoring in the bottom quartile, average range and top quartile in letter sound knowledge at pre-, mid- and post-test.
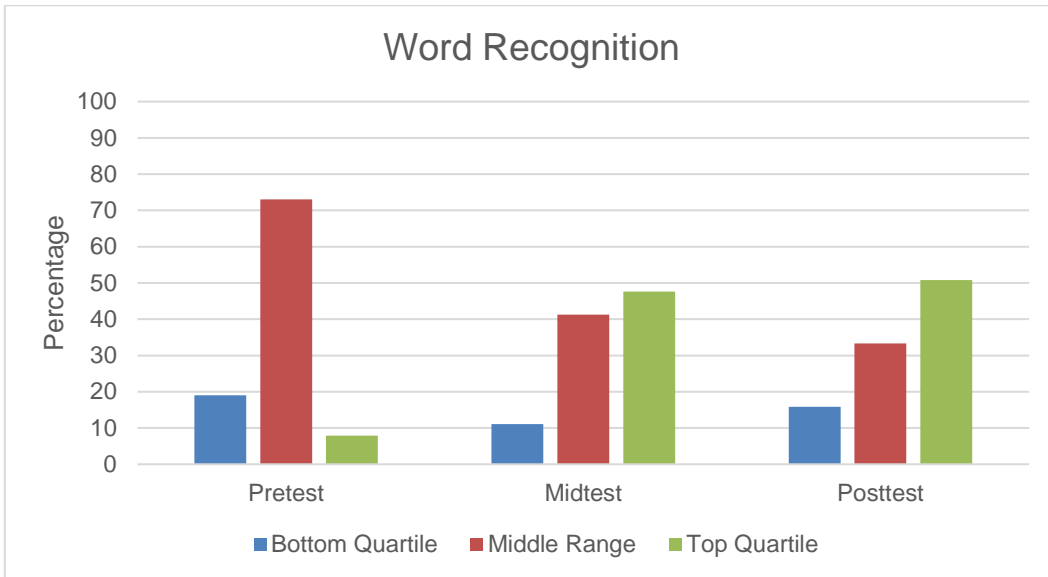
Figure 5. Percentages of students scoring in the bottom quartile, average range and top quartile in word recognition at pre-, mid- and post-test.
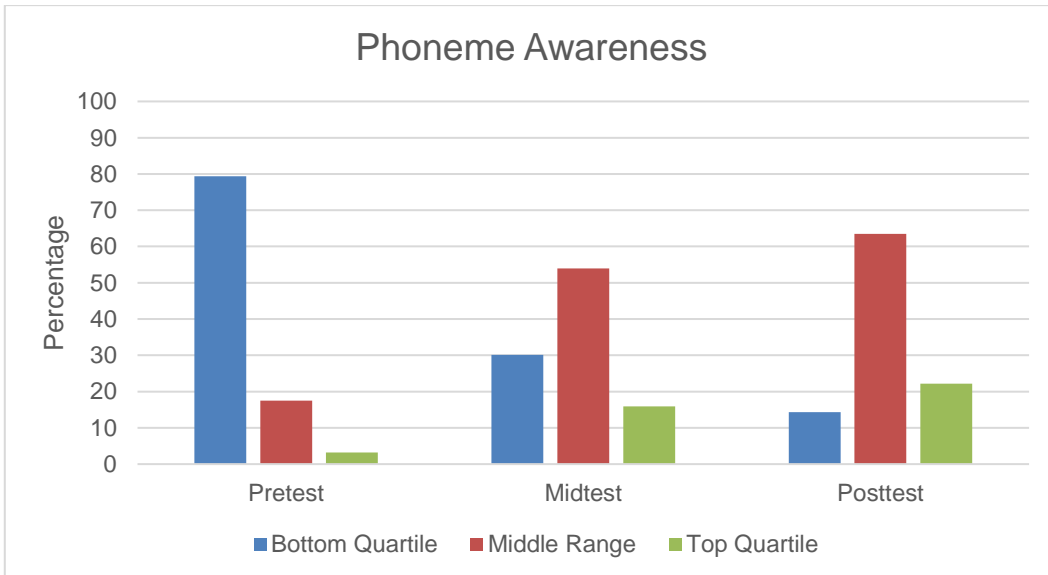

Figure 6. Percentages of students scoring in the bottom quartile, average range and top quartile in phoneme awareness at pre-, mid- and post-test.
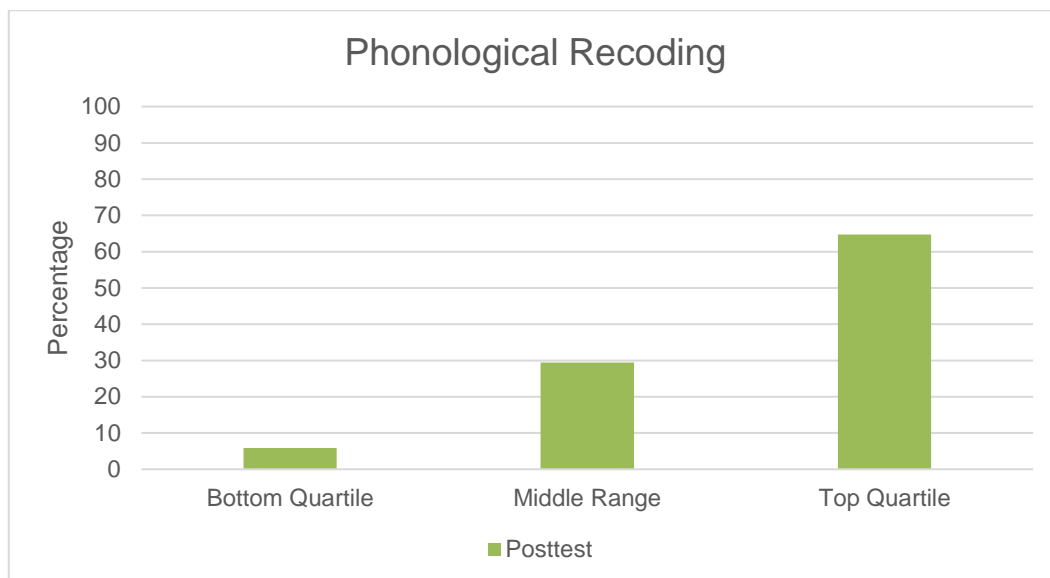
Figure 7. Percentages of students scoring in the bottom quartile, average range and top quartile in phonological recoding at post-test.

So, what may we conclude from these preliminary data? Recall that in the 2016 trial, the comparison classes in Sydney were receiving 'business as usual' which was a whole language approach. In Perth, the comparison class was receiving an exemplary explicit phonics program. It appears that instruction in InitiaLit is more effective than regular (essentially whole language) instruction (Sydney) and is at least as good as customised exemplary instruction (Perth), even when delivered by a relatively inexperienced teacher There is some evidence to suggest that the InitiaLit approach may lead to apparently slower growth at first but that the cumulative effect is such that it may lead to greater growth in the longer term.

The findings of the trial carried out in 2017 further support these conclusions, with students making progress over the year that was greater than the typical rate of progress in early reading skills. Students started the year below the average range of performance expected for their age and ended the year either within or above this average range. Significant and meaningful gains were not only seen over the first half of the year, when we would expect young children to make progress no matter the instruction provided but also over the second half of the year, indicating that the instruction provided was effective in developing these skills. Furthermore, there was a general shift of student out of the bottom quartile and middle range and into the top quartile, showing that the program helped students struggling with these skills while not limiting the students who were already performing at an average level.

References

Hulme, C., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., & Snowling, M. (2012). York assessment of reading for comprehension: Early reading test manual (Australian ed.). London: GL Assessment Ltd.

Martin, F., & Pratt, C. (2001). Martin and Pratt nonword reading test. Camberwell: Australian Council for Educational Research.

Neilson, R. (2016). Foundations of early literacy assessment (FELA). Jamberoo: Roslyn Neilson.

Wheldall, K., Wheldall, R., Madelaine, A., Reynolds, M., & Arakelian, S. (2016). 'What's new from MRU?': Recent research on reading instruction from the MultiLit Research Unit (MRU). Paper presented to the New South Wales Institute for Educational Research Conference on 'Literacy: What works and why', University of New South Wales, Sydney, November 15, 2016.