

MultiLit is committed to the promotion of scientific evidence-based best practice. The development of our programs is predicated on the best available scientific research evidence and we endeavour to provide evidence for the efficacy of our programs.

MacqLit was developed over a period of more than 10 years. Initially, a rough translation of the MultiLit Reading Tutor Program was developed for group instruction purposes and employed in our tutorial centres. Data on students' progress were collected during this period. Over time, the new group program evolved and again data were collected to show how students responded to the newer iterations. Over the years 2009–2015, a draft version of what has come to be known as MacqLit was developed and trialled and, following the program's official publication by MultiLit in 2016, more student data were collected in the period from 2016–2018.

Large-scale tutorial centre trials (2009–2014)

Students attending the Exodus Foundation tutorial centres in Sydney and Darwin in 2009–2011 were assessed before and after two terms of intensive literacy instruction that included an early version of the MacqLit program. It should be noted that these students received three hours of instruction daily and that while instruction was centred on the MultiLit group instruction program, other programs were also used, including, for example, Spelling Mastery. Students also participated in individual Reinforced Reading sessions with volunteers at the centres. Instruction was delivered by highly trained tutors employed by MultiLit. The literacy assessments were carried out by testers employed by MultiLit, but who were independent of the instructional delivery of the program.

A total of 362 students attended the programs at the various tutorial centres between 2009 and 2011 for whom data from pre- and post-testing were available. At program commencement, the average age of the students was 10 years, 5 months. Assessments were conducted before the start of the program and again at the end of their second term in the program, after about 19 weeks of instruction, on average.

At the start of the program, the average crude reading age for reading accuracy was 7 years, 5 months, while for reading comprehension it was 7 years, 1 month, as measured by the Neale Analysis of Reading Ability (NARA). Hence, students were three years below chronological age for reading accuracy and almost three-and-a-half years below chronological age for reading comprehension.

Based on analyses of raw scores, statistically significant gains ($p < .001$) were made on all of the six literacy measures used: reading accuracy (NARA), reading comprehension (NARA), single word reading (Burt Word Reading Test), word spelling (South Australian Spelling Test), passage reading fluency (Wheldall Assessment of Reading Passages or WARP), and phonological recoding (Martin and Pratt Nonword Reading Test). The effect sizes were large (Partial eta squared ≥ 0.8) for all measures.

On average, these students made crude gains of:

- ▶ 13 months in reading accuracy;
- ▶ 6 months in reading comprehension;
- ▶ 14 months in single word reading;
- ▶ 20 months in spelling;
- ▶ 20 months in phonological decoding, and;
- ▶ 35 (50%) more words read correctly per minute.

As stated, it should be emphasised that the results gained in this model were based on a three-hour intensive program that incorporated more reading and spelling practice than that offered in MacqLit. Although the program was very successful, it was acknowledged that schools would not be able to operate such a time-intensive model. For this reason, adaptations were made to the program.

In 2014, a more cost-effective model of program delivery was implemented in the tutorial centres. To reach more students in need of assistance, the number of students receiving instruction in the program was increased and program delivery was divided into two daily two-hour sessions, as opposed to only one. This was implemented with 20 students attending each session – one held each morning and one each afternoon – for five days a week. Students received group instruction in decoding, sight words, spelling, vocabulary and comprehension, along with group reading and individual Reinforced Reading with volunteers at the centres. The number of instructional hours was reduced from three to two hours per group. While most of the instruction involved an early version of MacqLit, additional programs such as Spelling Mastery were, again, still used at this point.

Full data were available for 74 students, who received the modified group instruction program for 19 weeks during Semester 2, 2014. At program commencement, the average age of the students was 10 years, 4 months. The average crude reading age at the start of the program for single-word reading, as measured by the Burt Word Reading Test, was 7 years, 9 months – more than two-and-a-half years below chronological age.

Again, based on analyses of raw scores, statistically significant gains ($p < .001$) were made on all measures: single word reading (Burt), spelling (South Australian Spelling Test), phonological decoding (Martin and Pratt), reading comprehension (Test of Everyday Reading Comprehension;) and passage reading fluency (WARP). The effect sizes were large (Partial eta squared > 0.8) for gains on all measures, ranging from 1.32 to 2.59. On average, these students subsequently made crude gains of:

- ▶ 12 months in single-word reading;
- ▶ 14 months in spelling;
- ▶ 18 months in phonological recoding, and;
- ▶ 22 (34%) more words read correctly per minute.

These gains achieved by students on all measures provided sound evidence of the efficacy of our two-hour group instruction model.

Large-scale school-based trials (2016–2018)

A more recent analysis was conducted with data collected between 2016 and 2018. Here, the focus was on 292 struggling readers in Years 3 to 6, who attended MacqLit programs at school for two terms (~19 weeks). Instruction was delivered by highly trained tutors employed by MultiLit. Assessments were administered by testers also employed by MultiLit, who were independent of the instructional delivery of the program.

Results showed that the group made substantial and statistically significant gains ($p < .001$) on raw score measures of word reading (Burt Word Reading Test), nonword reading (Martin & Pratt), spelling (South Australian Spelling Test) and passage reading fluency (Wheldall Assessment of Reading Passages).

The improvements were associated with large effect sizes (Partial eta squared > 0.5). Standardised scores were available for the nonword reading measure, and these scores indicated that the improvements were significantly ($p < .001$) beyond what would be expected for age-related change or regression to the mean.

A subset of the original 292 children ($n=111-181$) also participated in additional passage reading assessments, and again the sample showed substantial and statistically significant gains ($p < .001$) on raw score measures of:

- ▶ Passage reading accuracy ($n=111$ for NARA-3; $n=181$ for York Assessment of Reading Comprehension)
- ▶ Passage reading comprehension ($n=111$ for NARA-3; $n=181$ for York Assessment of Reading Comprehension)

The improvements were associated with large effect sizes (Partial eta squared >0.3). Standardised scores indicated that the improvements on both passage reading accuracy and comprehension measures were significantly beyond what would be expected for age-related change.

Evidence from randomised control trials

We have also completed, with our former doctoral student, Dr Jennifer Buckingham, a modest but rigorous randomised control trial (RCT) of an early version of MacqLit. In this RCT, instruction was delivered daily for one hour comprising exclusively MacqLit group program prototype materials: no other instructional materials or programs were used in the trial.

The trial was carried out in a school in a socially disadvantaged area of New South Wales. Students deemed eligible for the program (being in the bottom quartile for reading performance) were randomly allocated to experimental ('treatment') and control conditions. The students in the experimental treatment condition received the program in small groups for an hour each day for three school terms, while the control group continued to experience their usual classroom literacy activities. At the end of three terms, the conditions were reversed: the original experimental group now received regular classroom literacy instruction while the original control group became 'experimental group 2' and received the program in small groups. Instruction then continued for another three school terms.

Students were tested initially on a battery of tests of reading and related skills, and then retested again after three and six terms. Participant attrition occurred in both trials for a number of reasons, including their no longer being of an appropriate age for the program, or students leaving the school to attend high school. Attrition is a common occurrence within a demographic of this age.

The detailed findings from the two phases of the crossover design study are reported fully in two peer-reviewed academic journal articles (Buckingham, Beaman & Wheldall, 2012; Buckingham, Beaman-Wheldall & Wheldall, 2014) and will not be reported in detail here. This RCT initially comprised 30 students, with 26 students remaining for the second phase of the trial, equally distributed across the two conditions.

In the first phase of the study (three terms of instruction), the original treatment group (MacqLit) made greater gains than the control group and a large and statistically significant mean difference in gain between the groups was evident for the measure of phonological recoding (reading nonwords), with a large effect size (Partial eta squared = 0.94). Minimal effect sizes were apparent for the other literacy measures employed and they were not statistically significant.

In phase two of the study, however, the original experimental group did not make further significant gains on any measure but the former control group, now 'experimental group 2', made very large (and significant) gains on all measures, clearly demonstrating the efficacy of the program (Buckingham, Beaman-Wheldall and Wheldall, 2014) (see Figures 1 to 5 for details). Note that the teachers in this study were new to the program and initially inexperienced. We attribute the greater gains on all measures in the second phase of the study to their increased confidence and competence in the delivery of the program.

Efficacy data from trials conducted by researchers independent of MultiLit

At this point, there have been no studies involving MacqLit completed by other researchers. We would welcome other researchers to conduct efficacy studies on the program and will endeavour to facilitate their research.

Summary

As may be seen, the MacqLit small group instruction program has been shown consistently to deliver significant and substantial gains in literacy performance for older low-progress readers.

References to published efficacy studies

- Buckingham, J., Beaman, R., & Wheldall, K. (2012). A randomised control trial of a MultiLit small group intervention for older low-progress readers. *Effective Education, 4*, 1–26.
www.doi.org/10.1080/19415532.2012.722038.
- Buckingham, J., Beaman-Wheldall, R., & Wheldall, K. (2014). Evaluation of a two-phase experimental study of a small group ('MultiLit') reading intervention for older low-progress readers. *Cogent Education, 1*, 962786.
- Limbrick, L., Wheldall, K., & Madelaine, A. (2012). Do boys need different remedial reading instruction from girls? *Australian Journal of Learning Difficulties, 17*, 1–15.
- Wheldall, K. (2009). Effective instruction for socially disadvantaged low-progress readers: The Schoolwise program. *Australian Journal of Learning Difficulties, 14*, 151–170.
- Wheldall, K., Beaman, R., & Langstaff, E. (2010). 'Mind the gap': Effective literacy instruction for indigenous low-progress readers. *Australasian Journal of Special Education, 34*, 1–16.

Figure 1. Martin & Pratt Nonword Reading Test mean scores for Phase 1 experimental (E1) and Phase 2 experimental (E2) groups at pre-test, post-test 1 and post-test 2.

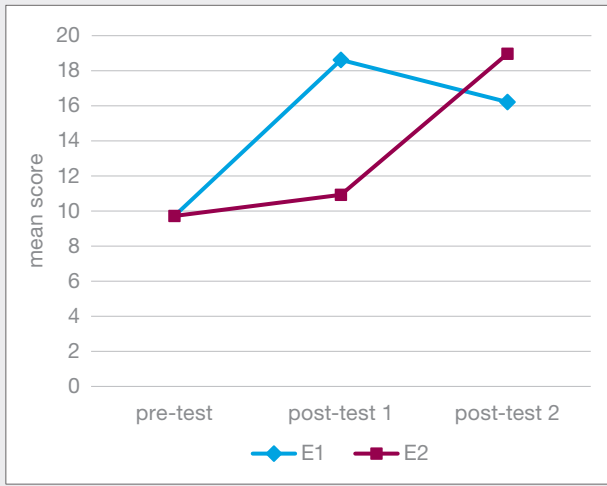


Figure 2. Burt Word Reading Test mean scores for Phase 1 experimental (E1) and Phase 2 experimental (E2) groups at pre-test, post-test 1 and post-test 2.

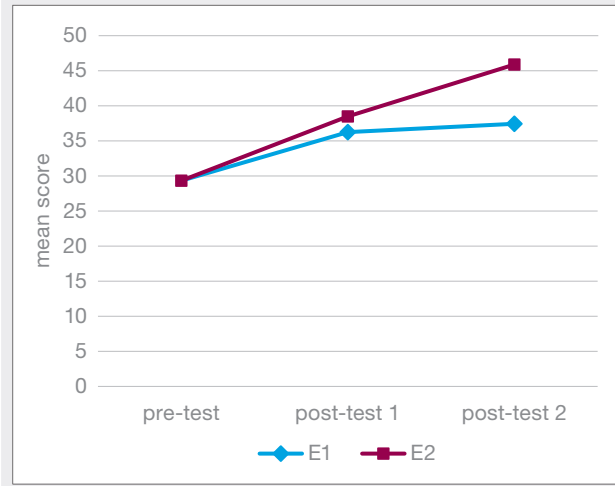


Figure 3. Wheldall Assessment of Reading Passages (WARP) mean scores for Phase 1 experimental (E1) and Phase 2 experimental (E2) groups at pre-test, post-test 1 and post-test 2.

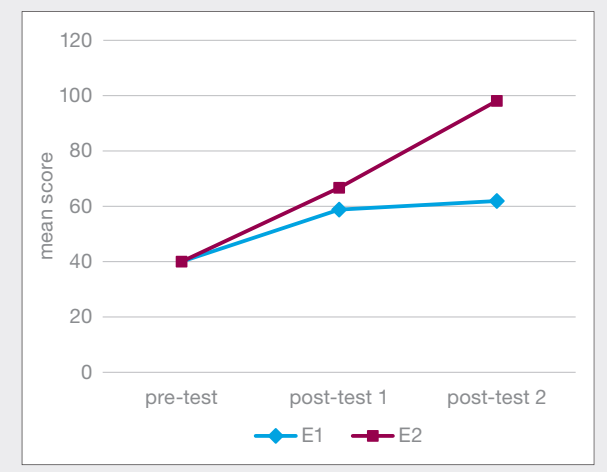


Figure 4. Neale Analysis of Reading Ability (Accuracy) mean scores for Phase 1 experimental (E1) and Phase 2 experimental (E2) groups at post-test 1 and post-test 2.



Figure 5. Neale Analysis of Reading Ability (Comprehension) mean scores for Phase 1 experimental (E1) and Phase 2 experimental (E2) groups at post-test 1 and post-test 2.

