



Does Rainbow Repeated Reading Add Value to an Intensive Literacy Intervention Program for Low-progress Readers? An experimental evaluation

KEVIN WHELDALL, *Macquarie University, Sydney, Australia*

ABSTRACT *Forty low-progress readers in Years 2 to 7 attending an intensive literacy intervention program, conducted on two separate sites, were randomly allocated to two treatment conditions in order to determine the effect of supplementing the program with commercially produced repeated reading materials (the Rainbow Reading Program). The control group experienced the regular form of the literacy program, which includes within it a form of repeated reading, while the experimental group experienced the supplemented literacy program incorporating repeated listening of tape-recorded versions of the text the student is required to read. All students were tested on two measures of reading performance immediately prior to and immediately following almost a term (9 weeks) of daily instruction. Analyses of covariance of post-test reading scores, in which pre-test scores were covaried, revealed no differences between the two groups on either of the two reading tests, indicating that the supplementary program did not demonstrate enhanced efficacy.*

Introduction

Repeated reading of short sections of text by young and low-progress readers is now commonly suggested as a means of increasing reading performance and research has repeatedly demonstrated that this method can be effective (Dowhower, 1994; Samuels, 1979, 1997). Typically, students are required to read the same short section of text, say, daily for a week, ideally timing their performance and graphing their progress. Samuels (1997), in 'Author notes' to the reprint of his seminal 1979 paper, comments that almost 200 studies have been completed since its original publication and draws several pertinent conclusions:

1. The original finding has been replicated; that is, a high degree of accuracy and speed develops on the practiced text. We can give accuracy and speed another name if we wish—fluency.
2. There is transfer of fluency to other portions of the text, even the parts that were not specifically practiced. Some have even reported improvements in comprehension.

3. Repeated reading is the most universally used remedial reading technique to help poor readers achieve reading skill.
4. Repeated reading is now widely used to teach reading in foreign languages. (pp. 380–381)

Repeated reading can be readily employed for classroom use with little or no additional materials being needed and with relatively minimal demands on teacher time other than time taken to monitor the child's progress and to assist in selecting appropriate level text. The present study aims to evaluate the relative efficacy of an elaborated version of repeated reading known as the Rainbow Reading Program.

The Rainbow Reading Program (Pluck, 1996) is a read-along taped reading program involving repeated readings. It consists of a series of books with accompanying audio tapes, a training video, a teacher's manual and an assessment manual. The books are colour coded according to the different reading levels. Each level includes both narrative and expository text appropriate for primary school aged students.

The program uses audio cassettes of an oral presentation of the text to support the students' reading of the text, thereby facilitating student comprehension and reading for meaning. By reading the same text repeatedly, the reader's fluency increases. Students request a 'conference' with the teacher when they believe that they have mastered the text, in order to move on to another book or another level.

The Teacher's Manual explains how to place the students at the starting level and when to move them onto the next text level. Accurate monitoring is facilitated by the provision of specially prepared recording sheets (Student/Teacher Conference, Student's Record of Practices, and Record of Books Read). The Rainbow Reading Program Assessment Manual provides all of the information needed for assessment of students using the Program which has been researched and trialed in and around Nelson, New Zealand.

In order to evaluate the program, it was included within the curriculum of the MULTILIT (Making Up Lost Time In Literacy) classrooms at Macquarie University Special Education Centre (MUSEC) for primary aged low-progress readers, complementing the existing intensive MULTILIT intensive literacy program. It was employed and evaluated informally during 1997 and for the first part of 1998. In the MULTILIT program there are eight children to one teacher, supported by at least one volunteer helper. The Rainbow Reading Program formed part of these students' daily work contracts.

The aim of this educational experiment was to determine the effect of adding the Rainbow Reading Program to the MULTILIT Program on the reading progress of low-progress readers. More specifically, the aim was to test whether the Rainbow Reading Program 'added value', in terms of enhanced program efficacy, to the MULTILIT Program which has already been reliably shown to yield substantial gains in reading progress over short periods of time. To this end, the Rainbow Program was added to the classroom activities of a randomly allocated sample of half of the new students attending the MULTILIT Program on two separate sites, the original MULTILIT site at MUSEC and in an outreach MULTILIT facility (known as 'Schoolwise') operating off-campus. The other half of the students, on both sites, continued to receive the standard MULTILIT Program which already incorporates a more informal operationalisation of repeated reading. Briefly, in this form of repeated reading students simply read the same passage (text at one level below their

TABLE I. Mean reading ages (and standard deviations) in months (Neale accuracy) for the experimental and control groups on the two sites ($n = 10$ per group/site)

Groups	Experimental ($n = 20$)	Control ($n = 20$)
Site A ($n = 20$)	85.7 (14.74)	85.3 (13.00)
Site B ($n = 20$)	95.3 (17.02)	93.8 (14.85)
Total ($n = 40$)	89.6 (14.27)	90.5 (16.26)

independent reading level) once daily for 1 minute over 4 consecutive days. The tutor marks their errors on a photocopy of the text and any errors made are discussed. The number of errors is subtracted from the total number of words read, and the number of words read correctly in 1 minute (wcpm) is recorded and graphed by the students.

Method

Sample and Design

The sample of participants in this study consisted of 40 low-progress readers from Years 2 to 7 commencing attendance at an intensive literacy program (MULTILIT) operating both on campus within Macquarie University Special Education Centre (MUSEC) Special School (site A) and in an outreach setting off-campus (site B). In order to access the program all students were assessed as being at least 2 years behind their age peers in terms of reading skills; 30% were girls and 70% boys.

Twenty students from each site were ranked (separately for each site) in terms of increasing reading accuracy, as measured by the Neale Analysis of Reading (Revised) (Neale, 1988), and were subsequently divided into ten pairs of students of comparable reading age for each site. One member of each pair was then randomly allocated to the experimental treatment (the addition of Rainbow Reading to their literacy program). The other members of the pairs acted as controls and received the standard form of the literacy intervention program (MULTILIT) (which included a less formally structured form of repeated reading). Thus, matched experimental and control groups each comprising ten low-progress readers, were established on each of the two sites (A and B). Students from Site A were drawn from Years 2 to 6 while students from Site B were drawn from Years 6 and 7. The mean (Neale accuracy) reading ages for the experimental and control groups in each site were very similar (see Table 1) while the reading ages for the students on Site B were, on average, about 9 months higher than for students on Site A.

Assessment Instruments and Test Procedures

As well as being assessed on the Neale Analysis of Reading—Revised for allocational purposes, all students were also assessed on two other measures of reading which constituted the dependent variables in this experiment, the Burt Word Reading Test and the Wheldall Assessment of Reading Passages.

The Neale Analysis of Reading—Revised (Neale, 1988). This widely used test measures, and provides reading ages for, reading accuracy (and also reading comprehension but this was not used in this study), based on the results of an Australian standardisation. A recent study by McKay (1996) has confirmed the Australian version of the Neale as an accurate and reliable measure of reading ability (p. 265).

The Burt Word Reading Test (New Zealand Council for Educational Research, 1981). The Burt test has a long history and is a measure of single-word recognition. The version employed was based on a standardisation carried out in New Zealand by the New Zealand Council for Educational Research. While doubts may be expressed over the utility of reading single words in isolation, it remains a robust test especially when used as part of a battery of reading tests.

The Wheldall Assessment of Reading Passages (WARP) (Wheldall, 1996). The experimental edition of the Wheldall Assessment of Reading Passages (WARP) (Wheldall, 1996) consists of a series of five 200-word passages, each passage comprising an entire story, specially written so that there could be no danger of students having encountered the passages before. Wheldall and Madelaine (1997) review the early development of the WARP but, briefly, the results of early investigations into the reliability and validity of the WARP have been favourable. The five passages have consistently correlated highly with each other (around 0.95) and, equally importantly, also yield means and standard deviations that are very similar to each other (Wheldall & Madelaine, 1997). Madelaine and Wheldall (1998) subsequently showed the test to have good criterion validity, correlating with Neale Analysis (accuracy) at over 0.8.

The tests were administered by trained and experienced research assistants. The Neale Analysis of Reading and the Burt Word Reading Test were administered according to the procedures outlined in their respective manuals (New Zealand Council for Educational Research, 1981; Neale, 1988:). Accuracy scores for the Neale and the score for the Burt were calculated and expressed as reading ages in months. Finally, the five passages of the WARP were administered to each student using the following procedure. Students were required to read the five passages out aloud after hearing the following instructions: 'I'd like you to read this passage as quickly and as carefully as you can. I'll start recording when you read the first word.'

A stopwatch was started when the student read the first word in the passage. The last word read at the end of 1 minute was marked and the student was asked to stop reading. Where appropriate the student was allowed to read to the end of the sentence or paragraph before stopping. If the student stalled on a word for more than 3 seconds, the word was supplied and scored as an error. If a word was omitted, one error was scored. If several consecutive words were omitted (e.g. a line was skipped), the total number of words omitted was subtracted from the total, and one additional error was scored. If a word was inserted, the inserted word was not counted, and one error was scored. If the word order was reversed, one error was scored. A numbered column on the tester's copy of each passage was used to assist in calculating the number of words read in 1 minute for each passage. From this score, errors were subtracted to calculate the number of words read *correctly* in 1 minute. The words read correctly per minute scores were averaged over the five passages to yield a single measure which served as the second dependent variable alongside the Burt.

Students were administered all three tests in the first week of Term 3 (July) and were retested on the Burt and the WARP in week 10, the final week of Term 3 (September), following approximately 9 weeks of intensive literacy instruction.

Analysis

As already stated, the aim of this study was to determine the effect of including Rainbow Reading within the MULTILIT intensive literacy intervention program to see whether this leads to significantly greater gains for low-progress readers than the 'standard' MULTILIT Program. Given a balanced, truly experimental design, incorporating pre- and post-test measures of the two dependent variables, the appropriate mode of analysis for this study is analysis of covariance (Lambourne & Wheldall, 1979; Dugard & Todman, 1995; Mok & Wheldall, 1995). As Mok and Wheldall (1995) argue:

... the analysis of covariance is to be recommended and, indeed, encouraged when analysing true experiments using pretest–posttest designs where subjects are randomly allocated to conditions. Increased sensitivity will result from incorporating the pretest scores as predictor variables into the analysis and any small initial differences between groups arising by chance will be taken care of. (p. 201)

Results

The average age of the sample was 137 months or nearly 11.5 years. From the results of the Neale Analysis (accuracy), this sample overall was over 47 months (nearly 4 years) behind their age peers for reading accuracy at the start of the study.

Neale Accuracy correlated highly with the Burt Word Reading Test at 0.88 and with the WARP at 0.9 (both $p < 0.001$), on the initial pre-intervention testings, and there was also a very high correlation between the WARP and the Burt for this sample (0.89, $p < 0.001$). The test/retest correlations for the Burt and the WARP, derived from scores obtained pre and post the intervention phase for the sample of 40 students as a whole, were also very high at 0.96 and 0.93 respectively ($p < 0.001$). This latter finding adds considerable confidence to the decision to covary pretest scores when analysing post-test scores following the experimental treatments.

The results of pre- and post-testing on the Burt and the WARP for the experimental and control groups are shown in Table 2. Analyses of covariance were performed on the post-test data for each reading measure separately, covarying pre-test scores, with group (experimental versus control) and site as factors, together with their interaction. For Burt scores, neither of the factors nor their interaction produced a significant effect and for WARP scores the only significant effect was for site ($F = 11.03$; $df = 1,35$; $p = 0.002$). Consequently, we may conclude that there was no evidence for statistically significant differences between the experimental and control groups and, hence, there is no support for the Rainbow Reading Program 'adding value' to the MULTILIT intensive literacy intervention program. The significant effect for site merely shows that for WARP scores, students on the MULTILIT site at MUSEC made significantly greater gains (mean gain 28 more words read correctly per minute) than students on the Schoolwise site (mean gain 15 more words read correctly per minute), an interesting finding but one which is not relevant to the present study.

TABLE II. Means (and standard deviations) for the Burt and WARP pre-test and post-test scores for the experimental and control groups

Tests		Experimental (<i>n</i> = 20)	Control (<i>n</i> = 20)
Burt	Pre-test	92.6 (16.30)	95.4 (18.34)
	Post-test	102.9 (20.16)	105.3 (23.89)
WARP	Pre-test	56.9 (35.13)	60.95 (34.30)
	Post-test	77.0 (36.03)	84.0 (37.15)

Over the term (9 weeks), the 40 students overall enrolled in the MULTILIT programs on the two sites made mean gains of 10.12 months ($SD = 7.34$) in terms of Burt reading age ($t, 38 = 8.72$) and 21.58 more words read correctly per minute ($SD = 13.62$) on the WARP ($t, 38 = 10.02$), both of which gains were highly statistically significant ($p < 0.0001$).

Discussion

The findings from this experiment provide no evidence for enhanced efficacy ('added value') of the MULTILIT intensive literacy intervention program by supplementing it with the Rainbow Reading Program which employs a taped version of repeated reading. This is, perhaps, not surprising, given the existing high efficacy of the MULTILIT program. In this study, for example, the mean gains for both Burt and WARP translate to impressive effect sizes of about 0.6. While noting the absence of a 'no-treatment' control by which to compare this effect size, it could be argued that it is asking too much of any program to demonstrate enhanced efficacy on top of such high existing efficacy.

To do so is not without precedent, however, as Wheldall *et al.* (1991) demonstrated in their study employing community volunteer tutors. Two randomly allocated groups of low-progress readers aged 9 to 12 years already attending an existing intensive literacy program were assigned to two tutoring conditions: trained tutoring employing the Pause, Prompt and Praise procedures (Wheldall & Mettem, 1985) versus a less structured form in which the tutor did little more than listen to the child read. Tutoring was given daily by community volunteers for approximately 30 15-minute sessions spread over a period of 7 weeks. Analyses of pre- and post-test measures revealed a statistically significant difference in reading accuracy gain favouring the Pause, Prompt and Praise group. The mean gains in months for reading accuracy were 13.5 months in the experimental group (using Pause, Prompt and Praise) compared with 7.8 months in the control group. Wheldall *et al.* argued that the control group gains were largely attributable to the pre-existing intensive literacy program and tentatively concluded that 'PPP tutoring appears to be a useful and powerful adjunct to an intensive skills-based literacy program' (p. 26).

Another objection that could be raised is that the MULTILIT teachers did not implement the program as rigorously or as thoroughly as the program requires. It must be conceded that no treatment integrity/procedural reliability data were taken on the implementation of the program and so no evidence can be provided to prove that the program was implemented properly. In response to this criticism, it can be argued that the MULTILIT teachers practised using the program for over a year

before the experiment was conducted and so they were very familiar with the materials. Moreover, they received intensive professional development and training from the author of the program, including practical demonstrations with similar students at MUSEC (but not those involved in the present study), immediately prior to the experimental evaluation commencing. All members of the MULTILIT teaching teams are highly qualified and experienced special education teachers specialising in literacy instruction for low-progress readers.

In fairness, however, it should also be emphasised that the Rainbow Reading Program supplementation did not adversely affect the efficacy of the MULTILIT program since no differences were apparent either way. More positively, it could be concluded that the Rainbow Reading Program was equally as effective as the existing, less formally structured, repeated reading section within the MULTILIT program which it replaced. To this extent, and the extent to which repeated reading generally contributes to the efficacy of the MULTILIT program (which is unknown at this stage), it can fairly be argued that Rainbow Reading offers a version of repeated reading which less experienced teachers might readily take up for use in their classrooms. Rainbow Reading was certainly much admired by the MULTILIT teaching team as a valuable addition to their literacy intervention materials and the students enjoyed using the Rainbow materials. Consequently, it will continue to be used in the program, particularly for those students who respond readily to the Rainbow taped-reading methodology. Given that it can be quite time-consuming to use, it will not, however, replace the existing repeated reading segment within the MULTILIT program for all students.

Finally, it should be emphasised that Rainbow Reading might well demonstrate its efficacy within more traditional classrooms where an intensive literacy intervention program is not in place. There is clearly a need for a further study assessing the efficacy of Rainbow Reading in such contexts.

Acknowledgments

I would like to thank Margrit Frischknecht and Robyn Beaman for their contributions to the research reported in this article; Anita Blows, Margaret Clayton, Freya Purnell and Fiona Wilkes for their assistance in data collection; and, again, Robyn Beaman, Margrit Frischknecht and Margaret Clayton for their helpful comments on earlier drafts of this article.

Correspondence: Professor Kevin Wheldall, Director, Macquarie University Special Education Centre, Macquarie University, Sydney, NSW 2109, Australia.

REFERENCES

- DOWHOWER, S.L. (1994) Repeated reading revisited: research into practice, *Reading and Writing Quarterly: Overcoming Learning Difficulties*, 10, pp. 343–358.
- DUGARD, P. & TODMAN, J. (1995) Analysis of pretest-posttest control group designs, *Educational Psychology*, 15, pp. 190–198.
- LAMBOURNE, R.D. & WHELDALL, K. (1979) The use of the analysis of covariance in educational research: panacea or pitfall for the unwary?, *Educational Studies*, 5, pp. 43–51.

- MADÉLAINE, A. & WHELDALL, K. (1998) Towards a curriculum-based passage reading test for monitoring the progress of low-progress readers using standardised passages: a validity study, *Educational Psychology*, 18, pp. 471–478.
- McKAY, M. (1996) The Neale analysis of reading ability revised—systematically biased?, *British Journal of Educational Psychology*, 66, pp. 259–266.
- MOK, M. & WHELDALL, K. (1995) Some reservations about the use of the analysis of covariance in educational research: a response to Dugard and Todman, *Educational Psychology*, 15, pp. 199–201.
- NEALE, M.D. (1988) *Neale analysis of reading ability—revised* (Hawthorn, Australian Council for Educational Research).
- NEW ZEALAND COUNCIL FOR EDUCATIONAL RESEARCH (1981) *Burt Word Reading Test: New Zealand revision* (Wellington, Lithoprint (NZ) Ltd).
- PLUCK, M.L. (1996) *The Rainbow Reading Program* (Nelson, Rainbow Reading Program Ltd).
- SAMUELS, S.J. (1979) The method of repeated readings, *Reading Teacher*, 32, pp. 403–408.
- SAMUELS, S.J. (1997) The method of repeated readings, *Reading Teacher*, 50, pp. 376–381.
- WHELDALL, K. (1996) *The Wheldall Assessment of Reading Passages: experimental edition* (Sydney, Macquarie University Special Education Centre).
- WHELDALL, K., COLMAR, S. & FREEMAN, L. (1991) Employing community volunteers to tutor low-progress readers using the Pause, Prompt and Praise tutoring procedures, *New South Wales Journal of Special Education*, 14, pp. 23–26.
- WHELDALL, K. & MADÉLAINE, A. (1997) Should we measure reading progress and if so how? Extrapolating the curriculum-based measurement model for monitoring low progress readers, *Special Education Perspectives*, 6(1), pp. 29–35.
- WHELDALL, K. & METTEM, P. (1985) Behavioural peer tutoring: training 16-year-old tutors to employ the 'pause, prompt and praise' method with 12-year-old remedial readers, *Educational Psychology*, 5, pp. 27–44.