Reading Recovery®: Does It Work?

by James W. Chapman and William E. Tunmer

Two recent reports on Reading Recovery® (RR), released by What Works Clearinghouse (WWC, 2007, 2008), claim that RR is an effective intervention even though only 5 out of 106 research papers met the organization's evidence standards (WWC, 2008). As Reynolds, Wheldall, and Madelaine (2009) noted, this number is small given the extent of the program's implementation and the amount of funding allocated to RR over 25 years.

Development of RR

Reading Recovery was developed in New Zealand by Marie Clay during the 1970s (Clay, 1979, 1980) and funded by the New Zealand Department of Education for adoption by schools throughout the country during the 1980s. The main aim of RR is to reduce substantially the incidence of reading failure by accelerating to average levels of performance the progress of 6-year-old children who show early signs of reading difficulty (normally children whose reading progress falls in the lowest 20% of the enrollment cohort in any given school). Students selected for RR receive 30 to 40 minutes of daily individual instruction over 12 to 20 weeks by a specially trained RR teacher. Emphasis is placed on developing within these children a self-extending system of reading strategies that involves the flexible use of multiple cues (syntactic, semantic, visual, graphophonic) to detect and correct errors while reading text. Decisions regarding exit, or "discontinuation," from RR are based on children reading at a level near their class average and attaining a reasonable degree of independence in reading.

Although developed for use in New Zealand there is little or no systematic New Zealand research showing that RR is effective, especially in relation to long-term benefits. This dearth of New Zealand research may seem surprising given that the RR program is used in other countries because of the belief that it has been successful in New Zealand. This view was formed largely because of Clay's research on RR (Clay, 1979, 1980), and the impression that the program must be successful because of its implementation throughout the country.

Clay's RR studies have been criticized because of significant design flaws, including a) no matched group of poor readers or a proper control group, b) inappropriate use of multiple t-tests for analyzing gain scores, c) inclusion of only those RR students who were considered successful rather than all RR students, d) failure to account for spurious regression-towards-the-mean effects, e) using only performance measures devised by Clay rather than independent standardized tests, and f) intervention and comparison groups not equivalent at baseline (Center, Wheldall, Freeman, Outhred, & McNaught, 1995; Nicholson, 1989).

Implementation of RR throughout New Zealand was achieved largely on the basis of Clay's own claims about the efficacy of the program, and because of support from the Department of Education. This official recognition was interpreted in other countries (e.g., United Kingdom) as evidence that the program must be effective (Soler & Openshaw, 2006).

Independent Studies of RR

Only three independent studies of RR have been conducted in New Zealand. The New Zealand Department (later renamed *Ministry*) of Education funded all three studies. Glynn, Crooks, Bethune, Ballard, and Smith (1989) found that the modest gains observed for most children successfully discontinued from RR largely disappeared within two years following completion of the program. Chapman, Tunmer, and Prochnow (2001) found that children selected for placement in RR and successfully discontinued from the program were on average 6 months behind their same-age peers at discontinuation, and 12 months below their same-age peers on standardized measures of reading performance 1 year following discontinuation. The discontinued RR children performed no better following their exit from the program than a group of poor readers who did not receive RR. Moreover, the RR children's rate of performance on a number of measures showed no acceleration effects during or after the RR program.

The third New Zealand study was undertaken by the New Zealand Council for Educational Research (NZCER) to investigate the effectiveness of RR in New Zealand (McDowall, Boyd & Hodgen, 2005). The design, however, precluded any scientifically based conclusions from being made because the data included national RR data returns for one year (2003) for RR children only, with no control group. Other data were based on perceptions of the effectiveness of RR obtained from responses to surveys and interviews with principals, RR teachers, the three New Zealand RR national trainers and seven RR tutors, information obtained from 30 focus groups that comprised RR teachers and that were led by their RR tutors, and information from eight "successful" RR schools selected by RR tutors. In short, there was no evidence demonstrating the *effectiveness* of RR with regard to either (1) achieving its goal of accelerating the reading progress of students experiencing early literacy learning difficulties, or (2) determining whether RR is the *most* effective approach for meeting the needs of struggling readers (Chapman, Greaney, & Tunmer, 2007).

New Zealand evidence regarding the effectiveness of RR is either seriously flawed (Clay, 1979, 1980; McDowall et al., 2005) or shows minimal to no gains as a result of placement in the program (Chapman et al., 2001; Glynn et al., 1989). Despite the lack of robust evidence in support of RR, it remains as a national Ministry of Education funded program.

Beyond New Zealand, several investigations and extensive reviews of the RR program have appeared in the literature (e.g., Center et al., 1995; Center, Freeman, & Robertson, 2001; Elbaum, Vaughn, Hughes, & Moody, 2000; Hiebert, 1994;

Continued on page 22

Morris, Tyner, & Perney, 2000; Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994; Shanahan & Barr, 1995; Wasik & Slavin, 1993). There is some convergent evidence that RR can be effective for some children, but RR has not been shown to be more effective than other, often less expensive, programs. Indeed, on the basis of a comprehensive and stringent meta-analysis of one-to-one tutoring programs in reading, Elbaum et al. (2000) concluded as follows:

> Overall, the findings of this meta-analysis do not provide support for the superiority of Reading Recovery over other one-to-one reading interventions. Typically, about 30 percent of students who begin Reading Recovery do not complete the program and do not perform significantly better than control students. As indicated in this meta-analysis, results reported for students who do complete the program may be inflated due to the selective attrition of students from some treatment groups and the use of measures that may bias the results in favor of Reading Recovery students. Thus it is particularly disturbing that sweeping endorsements of Reading Recovery still appear in the literature. (p. 617)

Given that RR involves intensive, one-to-one instruction, it should come as no surprise that some studies show the program to be an effective intervention for some children with reading difficulties (Reynolds & Wheldall, 2007). One-to-one instruction is much more effective than classroom instruction (Bloom, 1984). The important question, however, which has not been satisfactorily answered in favor of RR, is this: Holding the basic parameters of the RR program constant, namely, that it involves one-to-one instruction for 30 to 40 minutes per day for 12 to 20 weeks by a specially trained teacher and that it supplements regular classroom reading instruction, are the specific procedures and instructional strategies of RR more effective than any other one-to-one (or small group) tutoring program for struggling readers? The answer at present, must be "no." But an equally important question could well be in regard to whether RR can only be delivered by means of one-to-one instruction.

Theoretical Underpinnings of RR

Elbaum and colleagues (2000) noted that while there is a widespread belief that one-to-one instruction is more effective than instruction delivered to more than one student at a time, there is "little systematic evidence to support this belief" (p. 606). These authors presented a compelling argument in support of small group instruction: "Each additional student that can be accommodated in an instructional group represents a substantial reduction in the per-student cost of the intervention, or alternatively, a substantial increase in the number of students that can be served" (p. 606). Support for this contention comes from Iversen, Tunmer, and Chapman (2005) who reported that an early intervention program based on the RR format and including explicit, out of context teaching of word analysis skills could be developed for pairs of struggling readers that would allow them to make accelerated progress. Their study showed that although RR instruction in pairs required somewhat longer lessons (43 minutes versus 33 minutes), there were no major differences between children taught in pairs compared to those taught in one-to-one lessons on any measures at discontinuation. Children taught in pairs performed within the normal range on all measures and these positive effects were maintained on end-of-year measures. Thus, by increasing instructional time by about a quarter, RR teachers could double the number of students served without making any sacrifices in outcomes. Additional benefits were derived by supplementing the standard RR instructional approach with the teaching of explicit, out-of-context word analysis skills.

Explicit attention to the development of word analysis skills runs counter to the instructional philosophy of RR. Clay (1993) stressed the importance of encouraging children to use many cues while reading, constantly cross-checking one source of cues against another. She wrote that meaning is "the most important source of information," and that "the most important test for the child is "Does it make sense?" (1991, p. 292). Thus, the child uses word-level information primarily for confirmation of language predictions: "The child checks language predictions by looking at some letters" and "can hear the sounds in a word he speaks [i.e., predicts] and checks whether the expected letters are there" (Clay, 1993, p. 41). As Clay (1991) stated: "In efficient rapid word perception, the reader relies mostly on the sentence and its meaning and some selected features of the forms of words. Awareness of the sentence context (and often the general context of the text as a whole) and a glance at the word enables the reader to respond instantly" (p. 8). Clay (1993) specifically stated that children should be discouraged from relying too heavily on word-level cues: "If the child has a bias towards letter detail the teacher's prompts will be directed towards the message and the language structure" (p. 42). That is, when children show a *preference* for using word-level information to identify unknown words in text, Clay recommends that the teacher should divert their attention *away* from such information.

This text-based instructional emphasis reflects Clay's strong theoretical orientation and the instructional underpinning of RR, which was designed to complement the predominantly whole language approach to regular classroom instruction in New Zealand. In most New Zealand classrooms, students are taught what they need to know to learn to read *incidentally* (i.e., "as the need arises") through frequent encounters with interesting reading materials. According to Smith and Elley (1994), two leading proponents of the whole language approach in New Zealand, "Children learn to read themselves; direct teaching plays only a minor role" (p. 87).

Clay's instructional philosophy, reflected in the RR program, stresses the importance of using information from many sources in identifying unfamiliar words without recognizing that skills and strategies involving phonological information are of primary importance in beginning literacy development. The RR

view of literacy teaching, and the theoretical assumptions on which it is based, were rejected by the scientific community over the past two to three decades (e.g., Gough & Juel, 1991; Perfetti, 1992; Pressley, 2006; Spear-Swerling & Sternberg, 1996; Stanovich, 1991; Tunmer & Chapman, 2003). As Pressley (2006) pointed out, "the scientific evidence is simply overwhelming that letter-sound cues are more important in recognizing words…than either semantic or syntactic cues" (p. 21), and that "teaching children to decode by giving primacy to semantic-contextual and syntactic-contextual cues over graphemic-phonemic cues is equivalent to teaching them the way weak readers read!" (p. 164). The use of letter-sound relationships is the basic mechanism for acquiring word-specific (or sight word) knowledge, including knowledge of irregularly spelled words.

In support of these claims, Chapman and colleagues (2001) found in a longitudinal study of beginning literacy development in New Zealand that students selected by their schools for RR were, without exception, experiencing severe difficulties in detecting sound sequences in words (i.e., phonological awareness) and in relating letters to sounds (i.e., alphabetic coding) during the year preceding entry into the RR program. Participation in RR did not appreciably reduce these deficiencies, and the failure to remedy these problems severely limited the immediate and long-term effectiveness of the program. Progress in learning to read following participation in RR was strongly related to phonological processing skills at discontinuation from the program. Similar findings have been reported by Center and colleagues (1995) in Australia, and by Iversen and Tunmer (1993) in the United States.

Changing the goal of word-level instruction in RR from reading a specific text (with word analysis activities arising primarily from the student's responses during text reading) to learning skills and strategies that may generalize to all texts, does not mean adopting a rigid skill-and-drill approach in which word-level skills are largely taught in isolation with little or no connection to actual reading. Although struggling readers should receive explicit and systematic instruction in letter-sound patterns and word identification strategies outside the context of reading connected text, they should also be taught how and when to use this information during text reading through demonstration, modeling, direct explanation, and guided practice. It cannot be assumed that struggling readers who are successful in acquiring word analysis skills will automatically transfer them when attempting to read connected text.

In support of these claims, Iversen and Tunmer (1993) found that the effectiveness of RR could be improved considerably by incorporating more intensive and explicit instruction in phonological awareness and the use of letter-sound relationships (especially orthographic analogies), in combination with strategy training on how and when to use this knowledge to identify words while reading text and to spell words while writing messages. Morris and colleagues (2000) examined the effectiveness of Early Steps, a first grade reading intervention program that is very similar to RR, especially in the emphasis it places on contextual reading and writing. A fundamental difference, however, is that Early Steps also includes direct, systematic study of orthographic patterns that is "purposefully isolated from meaningful text so that the child can pay full attention to the patterns being studied" (p. 682). In discussing this important distinction between Early Steps and RR, Morris and colleagues argued that some children benefit from studying a single information source (e.g., spelling patterns) in isolation while simultaneously being offered the chance to integrate this knowledge in contextual reading and writing. With this important difference, they found that Early Steps was highly effective, especially for those children who were most at risk.

In general, there are three major advantages in providing beginning and struggling readers with explicit and systematic instruction in orthographic patterns and word identification strategies outside the context of reading connected text rather than relying solely on "mini-lessons" given in response to students' oral reading errors during text reading. First, instruction in word analysis skills that is deliberately separated from meaningful context allows students to give full attention to the letter-sound patterns that are being taught, as well as avoid having their text reading overly disrupted. Second, this instructional approach helps students to learn word-decoding skills that may be useful in reading all texts, not just a specific text. Third, including isolated word study in beginning and remedial reading programs helps to ensure that struggling readers see the importance of focusing on word-level cues to identify unfamiliar words in text rather than using context to supplement word-level information. As Pressley (2006) noted, one of the major distinguishing characteristics of struggling readers is their tendency to rely heavily on sentence context cues to compensate for their deficient alphabetic coding skills.

These considerations relate to the most serious shortcoming of RR, which is the differential effectiveness of the program. The program appears to be beneficial for some struggling readers but not others, as indicated by the high percentage (around 15% in New Zealand but up to 30% elsewhere) of RR students who do not complete the program but, instead, are "referred on" by their RR tutor for further assessment and possible additional remedial assistance (Chapman et al., 2007; Elbaum et al., 2000). There is also New Zealand evidence that a significant number of the lowest performing 6-year-olds are excluded from RR because they are considered "not ready" or "less likely" to benefit from the program than others, or are withdrawn early from the program because they failed to make "expected rates of progress" during the first few lessons (Chapman et al., 2007). Such a practice inflates the results. And, as noted previously, many successful (i.e., discontinued) RR students do not maintain the gains made in the program. Reynolds and Wheldall (2007) found from their review of research on RR that the program "has not demonstrated that it works for the students who are most at-risk for failing to learn to read" (p. 213), leading them to conclude that "the success of the program appears to be inversely related to the severity of the reading problem" (p. 209).

Conclusions

Changes in the instructional approach of RR have simply not kept pace with contemporary scientific research. As Church (2005) noted in a comprehensive review of research on accelerating reading development in low achieving children, RR

Continued on page 24

> …was designed in the 1970s prior to most of the modern research into how children learn to read. Not surprisingly, therefore, it lacks a number of elements which have been found by research to be essential in teaching low achieving children how to read. The most notable omissions are the failure to assess phonemic awareness, decoding fluency, or reading fluency, and the failure to provide systematic instruction and practice on phonemic awareness, decoding fluency and reading fluency to those students who are lacking in these kinds of skills. These are major shortcomings. (p.13)

The WWC reports fall short in conveying important information relating to the effectiveness of the RR program. Therefore, we have argued (Tunmer & Chapman, 2003) that fundamental changes to the program, based on contemporary research, should occur and are very likely to improve the effectiveness of the program, both in terms of outcomes and cost.

References

Bloom, B. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, *13*, 4–16.

Center, Y., Freeman, L., & Robertson, G. (2001). The relative effect of a code-oriented and a meaning-oriented early literacy program on regular and low progress Australian students in year 1 classrooms which implement Reading Recovery. *International Journal of Disability, Development and Education*, *48*, 207–232.

Center, Y., Wheldall, K., Freeman, L., Outhred, L., & McNaught, M. (1995). An evaluation of Reading Recovery. *Reading Research Quarterly*, *30*, 240–263.

Chapman, J. W., Greaney, K. T., & Tunmer, W. E. (2007). How well is Reading Recovery really working in New Zealand? *New Zealand Journal of Educational Studies*, *42*, 17–29.

Chapman, J. W., Tunmer, W. E., & Prochnow, J. E. (2001). Does success in the Reading Recovery program depend on developing proficiency in phonological-processing skills? A longitudinal study in a whole language instructional context. *Scientific Studies in Reading*, *5*, 141–175.

Church, J. (2005, December). *Accelerating reading development in low achieving children: A review of research*. Paper presented at the annual meeting of the New Zealand Association for Research in Education, Christchurch, New Zealand.

Clay, M. M. (1979). *The early detection of reading difficulties*. Auckland, New Zealand: Heinemann.

Clay, M. M. (1980). Reading Recovery: A follow-up study. *New Zealand Journal of Educational Studies*, *15*, 137–55.

Clay, M. M. (1991). *Becoming literate: The construction of inner control*. Auckland, New Zealand: Heinemann.

Clay, M. M. (1993). *Reading Recovery*. Auckland, New Zealand: Heinemann.

Elbaum, B., Vaughn, S., Hughes, M., & Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, *92*, 605–619.

Glynn, T., Crooks, T., Bethune, N., Ballard, K., & Smith, J. (1989). *Reading Recovery in context*. Wellington: New Zealand Department of Education.

Gough, P. B., & Juel, C. (1991). The first stages of word recognition. In L. Rieben & C. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 47–56). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hiebert, E. H. (1994). Reading Recovery in the United States: What difference does it make to an age cohort? *Educational Researcher*, *23*, 15–25.

Iversen, S., & Tunmer, W. E. (1993). Phonological processing skills and the Reading Recovery program. *Journal of Educational Psychology*, *85*, 112–126.

Iversen, S., Tunmer, W. E., & Chapman, J. W. (2005). The effects of varying group size on the Reading Recovery approach to preventive early intervention. *Journal of Learning Disabilities*, *38*, 456–172.

McDowall, S., Boyd, S., & Hodgen, E., with van Vliet, T. (2005). *Reading Recovery in New Zealand: Uptake, implementation, and outcomes, especially in relation to Maori and Pasifika students*. Wellington: New Zealand Council for Educational Research.

Morris, D., Tyner, B., & Perney, J. (2000). Early steps: Replicating the effects of a first-grade reading international program. *Journal of Educational Psychology*, *92*, 681–693.

Nicholson, T. (1989). A comment on Reading Recovery. *New Zealand Journal of Educational Studies*, *24*, 95–97.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high at-risk first graders. *Reading Research Quarterly, 29*, 8–39.

Pressley, M. (2006). *Reading instruction that works: the case for balanced teaching*. New York, NY: Guilford Press.

Reynolds, M., & Wheldall, K. (2007). Reading Recovery 20 years down the track: Looking forward, looking back. *International Journal of Disability, Development, and Education, 54*, 199–223.

Reynolds, M., Wheldall, K., & Madelaine, A. (2009). The devil is in the detail regarding the efficacy of Reading Recovery: A rejoinder to Schwartz, Hobsbaum, Briggs, and Scull. *International Journal of Disability, Development and Education, 56*, 17–35.

Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk readers. *Reading Research Quarterly, 30*, 958–996.

Smith, J. W. A., & Elley, W. B. (1994). *Learning to read in New Zealand*. Auckland, New Zealand: Longman Paul.

Soler, J., & Openshaw, R. (2006) *Literacy crises and reading policy: Children still can't read*. London, England and New York, NY: Routledge.

Spear-Swerling, L., & Sternberg, R. J. (1996). *Off track: When poor readers become "learning disabled."* Boulder, CO: Westview Press.

Stanovich, K. E. (1991). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly, 26*, 7–29.

Tunmer, W. E., & Chapman, J. W. (2003). The Reading Recovery approach to preventive early intervention: As good as it gets? *Reading Psychology, 24*, 337–360.

Wasik, B., & Slavin, R. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28*, 179–200.

What Works Clearinghouse. (2007). *Reading Recovery.* Retrieved from the Institute of Education Sciences, U.S. Department of Education website: http://ies.ed.gov/ncee/wwc/pdf/WWC_Reading_Recovery_031907.pdf

What Works Clearinghouse. (2008). *Reading Recovery intervention report*. Retrieved from the Institute of Education Sciences, U.S. Department of Education website: http://ies.ed.gov/ncee/wwc/pdf/wwc_reading_recovery_120208.pdf

***James Chapman, Ph.D.**, is professor of educational psychology and Pro Vice-Chancellor of the College of Education at Massey University. He has a Ph.D. in educational psychology from the University of Alberta, Canada. His research activities have focused on motivational aspects of learning difficulties and more recently, on factors associated with the acquisition of reading and the emergence of reading difficulties/dyslexia.*

***Bill Tunmer, Ph.D.**, is a Distinguished Professor of Educational Psychology in the College of Education at Massey University, New Zealand, and co-author of the seminal* Simple View of Reading. *His Ph.D. is from the University of Texas at Austin. His research activities are in the acquisition of reading and the emergence of reading difficulties/dyslexia.*