Reading Recovery (RR) was developed in New Zealand by Mary Clay in 1976 (Clay, 1985). This program was developed during the Whole Language era. Reading Recovery specifically caters to pupils in grade 1. Within the context of RR, children at risk in grade 1 are seen on an individual basis, 30 minutes a day for 12 to 20 weeks. The activities include letter recognition, rereading short graded books, reading new short graded books, writing words and sentences, decomposition and recomposition of words in writing, and reading sentences and short texts written by the pupil (Clay, 1985; Legault, 2006; Shanahan and Barr, 1995; Elbaum et al., 2000). Working on phonics (decoding) is not systematic and is only done when the child has difficulty (as it is prescribed in Whole Language).

The Reading Recovery program has several serious problems. Shanahan and Barr (1995) concluded in their research that RR obtains positive results, but that these results must be tempered since 10% to 30% of the pupils are withdrawn from the program, and thus from the final research results if they do not progress sufficiently, or if they are frequently absent. In their meta-analysis, Elbaum et al. (2000) conclude that the assertions of the RR program are unfounded since approximately 30% of struggling pupils were removed from the program and consequently from the final research results. A more recent meta-analysis by D’Agostino and Murphy (2004) concluded that RR is effective, but strangely failed to mention the exclusion of the samples mentioned by Shanahan and Barr (1995) and Elbaum et al. (2000). Excluding 30% of the weakest pupils is not a “small” flaw, it is a fundamental methodological deficiency. If, while researching the treatment of juvenile leukemia, we sys-

11. In fact, the Reading Recovery program is often extended into grade 2.
tematically excluded from the sample pool the subjects who died and those not responding to treatment, the medical world’s knowledge of this terrible disease would not be nearly as advanced as it is, and it would not have reached a success rate of 80% for this type of cancer. Thank God, researchers in the field of education are not the same as those researching AIDS or breast cancer!

Further examination shows another serious flaw, in my humble opinion, which has yet to be discussed in papers on RR; the evaluation methods used to indicate RR pupils’ progress. The majority of the methods employed are both simple and peripheral to the skill of reading.

The methods of evaluation generally used to validate the success of RR at the end of an intervention or school year are: 1) evaluating the recognition of upper and lower case letters of the alphabet; 2) reading three lists composed of 15 isolated words most frequently used in writing; 3) evaluating concepts of writing (upper and lower case, word, sentence, beginning of a book, punctuation…); 4) asking the pupil to write as many words as possible in 10 minutes; 5) dictating words for which every sound correctly identified is scored; 6) reading short graded books orally, where the level is determined by the last book read with an accuracy rate of 90% (Clay 1985; Elbaum et al., 2000; Legault, 2006; Shanahan and Barr 1995).

Tests 1, 2 and 3 are only relevant during the first weeks of learning to read. After this, these tests become inappropriate. There is a risk of over-evaluating RR pupils using these tests. For example, with test 1 (letters of the alphabet), a good reader at the grade 1 level can forget both the name and sound of lower case and upper case letters J, Q, Y and X, and can confuse B and P sounds without having any serious impact on his reading and comprehension. Nevertheless, an RR pupil can do better than a good reader on test 1 simply because he is still working
on letter recognition during RR sessions. Test 2 (reading a list of frequently used words) is based on word recognition (see Appendix A, p. 73), which is irrelevant when pupils know how to read. It is true that a good reader will normally do well in this type of evaluation, but a weaker pupil can do equally well since the reading evaluation material and writing activities during the RR sessions are composed of words most often used in writing (Clay, 1985; Legault, 2006). Test 3, based on writing concepts, is the RR test furthest from the skill of reading. This method of evaluation can be justified, but only when a student is just beginning to learn to read. This test will produce the same biased results as tests 1 and 2, where a struggling first grade reader can obtain a higher success rate than a strong reader simply because RR instruction emphasizes these test skills. This evaluation does not reveal the level of reading skill. Tests 4 and 5 are writing evaluations, which allow an indirect evaluation of decoding, but are weak measures of students’ ability to read. Once again, pupils in the RR program can have higher success rates, but these results are not accurate indicators of their reading skills. In summary, tests 1 to 5 emphasize peripheral aspects of reading, and some tests (1, 2 and 3) are too simple to successfully measure a student’s reading ability. Therefore, these tests do not demonstrate an accurate and relevant measure of the success rate of remedial reading intervention.

Only test 6 is directly linked to the skill of reading, but only measures the accuracy rate (see Appendix A, p. 73) of oral reading. The accuracy rate of oral reading is an essential aspect but if not accompanied by a measure of reading rate (see Appendix A, p. 73), the accuracy rate is meaningless. It is obvious that a child reading 50 words per minute with an accuracy rate of 95% has a far better level of oral reading than a child reading 25 words per minute with the same accuracy rate of 95%. Despite both pupils’ accuracy rate of 95%, it is likely the child with the better
reading rate will display a superior level of information retention and will be more likely to reason easily around the text read. Although reading rate and accuracy rates are essential elements of oral reading, one without the other presents inaccurate and misleading information.

The first five assessments used to evaluate and demonstrate progress in the RR program can provide information and insight at the start of a child’s learning to read, but has absolutely no significance as summative assessments when it comes to the skill of reading. Worse still, these assessments can produce inflated results for the RR pupils in comparison to pupils of the same age that have no difficulty reading.

Continuing with the topic of achievement assessments, RR has yet another major defect; it does not measure reading comprehension. A special education program that has been taught for 30 years and claims to rehabilitate poor readers, but fails to evaluate literal and inferential reading comprehension (see Appendix A, p. 73), is as unacceptable as someone evaluating an individual’s ability to snowboard by simply measuring whether he or she can put on boots and a helmet. Snowboarding means being able to descend a slope, and for those who are young and foolish, make daring jumps. Reading means being able to read texts while assimilating literal information and producing inferences, regardless of a child’s grade level.12

Finally, the last flaw of the RR program is economic. RR is very expensive; special education teachers in this program can only work with an average of 16 first grade pupils per year (Shanahan and Barr, 1995; Elbaum et al., 2000). Most educational systems cannot afford to maintain this program long term, unless they are in a wealthy country during an economic boom.

12. Legault (2006) concluded that RR is efficient in a study conducted in Ontario (Canada) using more or less the same assessment tests listed above, without testing reading comprehension. Not a very convincing doctorate thesis.
Despite its popularity in the English world, Reading Recovery is a free flow intervention program with many significant weaknesses. Elimination of up to 30% of weaker pupils in its programs, the use of simple and non-relevant tests, the lack of reading comprehension evaluation, and an inability to reach a larger number of pupils during a school year make the RR special education program worth avoiding.

To conclude, in light of the information provided there appears to be sufficient evidence to make one seriously doubt the effectiveness of special education programs. How can we explain the failure of special education interventions? Obviously, there are many concomitant factors. In the following pages, I call your attention to one underestimated factor: time and intensity, which is an important element of the DIR program.

**Time and Numbers**

Mathematical arguments are not valued in Québec education.\(^{13}\) Yet, numbers help us see, think, and understand more

---

\(^{13}\) This allergy to numbers and statistics is the reflection of, on the one hand, poor scientific knowledge in education and, on the other hand, the depreciation of rigour in the last decades in the world of education. It must be said that the dominating pedagogical trend for more than 30 years in Québec, one of romantic and highly ideological flavour, has not been compatible with rigour. The Ministry of Education of Québec and school boards do not have any interest in knowing how many grade 1 pupils have been recuperated by special education interventions, or how many grade 1 pupils know how to read in June, or the real impact the last educational reform is having on pupils’ learning. When they do question, they do so in such a light and non-rigorous way that one wonders if they are truly interested in knowing the answers. The chosen “study” methods are often surveys, group discussions or “expert” consultations. At the end of the “study”, conclusions drawn are expected answers that never contradict the mainstream ideological educa-
Bibliography


National Reading Panel. Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and its Implications for Reading Instruction. Rockville: National Institute of Child Health and Human Development, 2000.


Appendix H  Reasoning Reading  89
Appendix I  Over-Reading  93
Appendix J  Activity List  96
Appendix K  Evaluating Oral Reading  97
Appendix L  Guidelines to Consider When Evaluating Reading Reasoning and Comprehension  99
Appendix M  Questionnaire Samples  101
Appendix N  Test Samples  104

Bibliography  108