Response to Intervention Models: The Role of Benchmarks and Curriculum-based Measurement in Making Instructional Decisions

Alison Madelaine, Kevin Wheldall and Meree Reynolds

Presented by Alison Madelaine
Overview

• What is Response to Intervention (RtI)?
• How can RtI be implemented?
• What are some of the advantages and disadvantages of RtI?
• What is the role of benchmarks and curriculum-based measurement in making instructional decisions within RtI models?
What is RTI?

• An approach to service delivery in schools
  – early intervention/prevention

• An alternative method of identifying students with learning disabilities.
How can RTI be implemented?

RTI is implemented using tiers of more intensive instruction.

2? 3? 4 tiers?
Tier 1

• Quality, scientific, research-based reading instruction for all children (100%).
• There is still some debate about what this should include.
• Responsibility of general (regular) education
• Primary prevention
Tier 1

• Example: Reading instruction delivered to whole classes in a regular setting. Instruction would be direct, systematic, and include phonemic awareness, phonics, fluency, text comprehension and vocabulary.
Tier 2

- Careful screening to identify struggling readers
- Supplementary small group (3-4 students) instruction
- For, say, bottom 25% of children
- Direct, systematic, explicit, scientific evidence-based instruction
- Instruction can be delivered by teachers or teacher aides (under supervision).
Tier 2

• Example: Small group instruction, 3-4 times per week for 10-20 weeks.
Tier 3

- Careful monitoring to identify ‘stragglers’
- More intensive intervention (tertiary intervention)
- Individual, one-to-one instruction for perhaps >5%
- Direct, systematic, explicit, scientific evidence-based instruction
- More frequent progress monitoring
- More time devoted to instruction
Tier 3

• Example: 1:1 instruction with a special educator

• Sometimes involves very small groups (1-3 students)
The RtI model

- Tier 3 5% max
- Tier 2 (+3) 25%
- Tier 1 (+2+3) 100%
How do we determine responsiveness?

• Tier 1 ---> Tier 2

Is a one off screening measure enough?
How do we determine responsiveness?

Different Methods:
1. Median split
2. Normalisation
3. Benchmarks
4. Dual discrepancy
5. Slope discrepancy
How do we determine responsiveness?

• Do these methods identify the same non-responders?

• Some research has found that the different methods resulted in different students being identified as ‘nonresponders’.
How do we determine responsiveness?

Level of performance

AND

Slope of improvement (progress) [CBM]

* Both need to be more than 1 SD below peers.
How do we determine responsiveness?

• What happens if students improve (ie they become ‘responsive to intervention’)?
Advantages of RtI

• Eliminates a ‘wait to fail’ situation
• Improved student outcomes
• Ensures that services are provided based on educational need rather than labels or categories
• Better information about student progress
Advantages of RtI

• Better quality instruction
• Cost effectiveness
• Educational accountability
• Better integrated assessment and instruction
Disadvantages of RtI

- Costly (time and resources)
- Application can be inefficient
- Difficult to evaluate (lack of scientific research).
- Professional skills issues.
Disadvantages of RtI

• Monitoring issues.
• Difficulty defining ‘responsiveness’
• RtI models cannot operate successfully without adequate Tier 1 instruction.
Establishing provisional benchmarks for identifying young low-progress readers in Years 1 and 2

- Benchmarks = expectations of average progress
- Provide objective information
- Make it easier to determine priorities for intervention
- The use of benchmarks can overcome problems of inequity within an education system
Benchmarks for what?

- The benchmarks for literacy need to relate to crucial components of the early reading process.
- Reviews of research have established that the key components of an early reading program are:
  - Phonemic awareness
  - Phonics
  - Fluency
  - Vocabulary
  - Comprehension
Aim

The aim of the study was to establish preliminary performance benchmarks to identify the bottom 25 per cent of Year 1 and Year 2 students in NSW schools at two key points in the school year.
The study

• The participants were students from grades one and two (n=335) at two primary schools (162 Year 1 students, 173 Year 2 students)
• The schools had been nominated as performing at state average levels
• Passive consent procedures were used
• Students were assessed on a battery of tests assessing various aspects of reading and related skills
• Assessments were carried out at the beginning and in the middle of the school year
Assessment Tools

• The Burt Word Reading Test
• The Martin & Pratt Non-Word Reading Test
• The Wheldall Assessment of Word Reading Lists (WARL)
• The Wheldall Assessment of Reading Passages (WARP)
• The Peabody Picture Vocabulary Test (PPVT-IV)
• Sutherland Phonological Awareness Test (SPAT-R)
• The South Australian Spelling Test
Determining the Benchmarks

We used the raw scores for each assessment tool

• To identify “struggling” students we found the cut-off points for the lowest quartile of students at each testing point

(Students whose scores fall below these cut-off points at the corresponding point in their schooling would be considered to be “struggling” students in the areas tested by the instrument.)

• To determine performance goals for intervention we used estimates of average performance (40th percentile) at each testing point

(Students whose scores are at or above these goals at the corresponding point in their schooling would be considered to be achieving at average or better levels in the areas tested by the instrument.)
Benchmarks for the Burt Word Reading Test

<table>
<thead>
<tr>
<th>Literacy Variable (raw scores)</th>
<th>Beginning Year 1 (raw score)</th>
<th>Middle Year 1 (raw score)</th>
<th>Beginning Year 2 (raw score)</th>
<th>Middle Year 2 (raw score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BURT</td>
<td>40%</td>
<td>23</td>
<td>30</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>18</td>
<td>27</td>
<td>30</td>
</tr>
</tbody>
</table>

MultiLit Research Unit
## Benchmarks for the Martin & Pratt Non-Word Reading Test

<table>
<thead>
<tr>
<th>Literacy Variable (raw scores)</th>
<th>Beginning Year 1 (raw score)</th>
<th>Middle Year 1 (raw score)</th>
<th>Beginning Year 2 (raw score)</th>
<th>Middle Year 2 (raw score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martin &amp; Pratt</td>
<td>40%</td>
<td>7</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>4</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>
Benchmarks for the WARL

<table>
<thead>
<tr>
<th>Literacy Variable</th>
<th>Beginning Year 1 (raw score)</th>
<th>Middle Year 1 (raw score)</th>
<th>Beginning Year 2 (raw score)</th>
<th>Middle Year 2 (raw score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WARL</td>
<td>40%</td>
<td>26</td>
<td>46</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>18</td>
<td>39</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>54</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>46</td>
<td>58</td>
</tr>
</tbody>
</table>
## Benchmarks for the WARP

<table>
<thead>
<tr>
<th>Literacy Variable</th>
<th>Beginning Year 1</th>
<th>Middle Year 1</th>
<th>Beginning Year 2 (raw score)</th>
<th>Middle Year 2 (raw score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WARP</td>
<td>40%</td>
<td>-</td>
<td>51</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>-</td>
<td>39</td>
<td>62</td>
</tr>
</tbody>
</table>
Benchmarks for the Peabody Picture Vocabulary Test

<table>
<thead>
<tr>
<th>Literacy Variable</th>
<th>Beginning Year 1 (raw score)</th>
<th>Middle Year 1 (raw score)</th>
<th>Beginning Year 2 (raw score)</th>
<th>Middle Year 2 (raw score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT</td>
<td>40% 90</td>
<td>105</td>
<td>108</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>25% 84</td>
<td>96</td>
<td>103</td>
<td>109</td>
</tr>
</tbody>
</table>
## Benchmarks for the SPAT-R

<table>
<thead>
<tr>
<th>Literacy Variable</th>
<th>Beginning Year 1 (raw score)</th>
<th>Middle Year 1 (raw score)</th>
<th>Beginning Year 2 (raw score)</th>
<th>Middle Year 2 (raw score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPAT-R 40%</td>
<td>28</td>
<td>36</td>
<td>37</td>
<td>41</td>
</tr>
<tr>
<td>SPAT-R 25%</td>
<td>23</td>
<td>29</td>
<td>33</td>
<td>36</td>
</tr>
</tbody>
</table>
## Benchmarks for the South Australian Spelling Test

<table>
<thead>
<tr>
<th>Literacy Variable</th>
<th>Beginning Year 1 (raw score)</th>
<th>Middle Year 1 (raw score)</th>
<th>Beginning Year 2 (raw score)</th>
<th>Middle Year 2 (raw score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAS</td>
<td>40%</td>
<td>15</td>
<td>21</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>11</td>
<td>18</td>
<td>20</td>
</tr>
</tbody>
</table>

MultiLit Research Unit
Reference

The role of curriculum-based measurement of reading within the Response to Intervention model
What is curriculum-based measurement?

- What is curriculum-based assessment?
- How is CBM different?
- CBM of reading
  - Oral reading fluency
  - Word identification fluency
CBM within RtI

Assessment within a RtI model may include:

- Standardised tests
- Curriculum-based assessment
  - Curriculum-based measurement
CBM within RtI: Assessment functions

1. Screening
2. Progress Monitoring
3. Instructional Decision-making
Screening

CBM as a screening measure:
• To identify which students are at risk of not responding to tier 1 instruction
Progress Monitoring

CBM as a progress monitoring tool:

- To track the weekly progress of students at risk in Tier 1 instruction
- To track the weekly progress of students in Tier 2 (and Tier 3) instruction
CBM within RtI

Tier 1 instruction
• All students assessed less often (screening)
• At-risk students assessed weekly
• Low-progress readers identified for Tier 2 instruction

Tier 2 instruction
• All students assessed weekly
• Students making progress continue with Tier 2 instruction, and when able, move back to Tier 1 instruction (with continued monitoring)
• Students not making progress identified for Tier 3 instruction
What is the WARP?

• Wheldall Assessment of Reading Passages
• A measure of oral reading fluency
• Comprises a series of 200 word passage reading tests of similar difficulty
• Can be used as a screening tool or a progress monitoring tool within a RtI model
• 3 initial assessment passages and 10 progress monitoring passages
Last week Dad decided that buying things cost too much money. “We could make all sorts of things if I had a workshop,” he said. “Just a little one with a workbench and a few tools.”

Mum looked worried. She did not think it was a very good idea. “Great idea, darls. The only problem is that good tools cost a lot of money.” And Mum smiled. She thought that would be the end of Dad’s idea but she was wrong.

“No problem, love,” said Dad. “My friend has a shed and some tools that he wants to sell. He’ll let me have the lot for just three hundred dollars.”

Mum gulped. What could she try next?

“Why not go and get the shed and the tools right now,” she said. “Then you could fix the toaster and the dryer. And then why not look at the back gate. It keeps sticking.”

Mum went on and on. Dad looked a bit sick and then he gave in. “Perhaps not after all, darls. There’s a really good show on the television tonight. Maybe next week.”

Mum smiled. She knew that she would not hear about the shed or the tools again.
The WARP: Reliability

- Alternate Forms Reliability: 0.97-0.98
- Test-retest Reliability: 0.93-0.97
The WARP: Validity

Criterion Validity:

- Neale Accuracy: 0.87
- BURT: 0.85
- MultiLit Word Attack Skills Placement Test: 0.78
- TOWRE:
  - Sight Words: 0.89-0.91
  - Phonemic Decoding: 0.81-0.85
- NSW BST Literacy overall:
  - Year 3: 0.85
  - Year 5: 0.66
The WARP: Grade-based norms

<table>
<thead>
<tr>
<th>Grade</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Score</td>
<td>37</td>
<td>78</td>
<td>93</td>
<td>121</td>
<td>130</td>
</tr>
<tr>
<td>50% Range</td>
<td>16-56</td>
<td>45-104</td>
<td>66-111</td>
<td>94-155</td>
<td>107-154</td>
</tr>
<tr>
<td>Limits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: values below the lower figure indicate levels achieved by the bottom 25%; values above the upper figure indicate levels achieved by the top 25%.*
The WARP: Grade-based norms

<table>
<thead>
<tr>
<th>Grade</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term 1</td>
<td>57</td>
<td>86</td>
<td>107</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>Term 2</td>
<td>68</td>
<td>90</td>
<td>114</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>Term 3</td>
<td>37</td>
<td>78</td>
<td>93</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td>Term 4</td>
<td>47</td>
<td>82</td>
<td>100</td>
<td>123</td>
<td>130</td>
</tr>
</tbody>
</table>
The WARP: Bottom 25%
The WARP – Progress monitoring of individual students

Year 4 Low-progress Reader
The WARP – Progress monitoring of individual students

Year 4 Low-Progress Reader

WCPM

Pretest
Week 1
Week 2
Week 3
Week 4
Week 5
Week 6
Week 7
Week 8
Posttest
Week 9
Week
Week
Week
Week
Week
Week
Week
Post
What is the WARL?

- Wheldall Assessment of Reading Lists
- A measure of Word Identification Fluency
- Words randomly chosen from a database of frequently occurring words
- Appropriate for student who cannot yet read a passage of text
- Can be used as a screening tool or a progress monitoring tool within a RtI model
- Comprises a series of 3 Initial Assessment Lists and 10 Progress Monitoring Lists of 100 words each
### List a

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>I</em></td>
<td><em>and</em></td>
<td><em>said</em></td>
<td><em>to</em></td>
<td><em>he</em></td>
</tr>
<tr>
<td><em>you</em></td>
<td><em>that</em></td>
<td><em>they</em></td>
<td><em>on</em></td>
<td><em>for</em></td>
</tr>
<tr>
<td><em>but</em></td>
<td><em>up</em></td>
<td><em>at</em></td>
<td><em>we</em></td>
<td><em>can</em></td>
</tr>
<tr>
<td><em>not</em></td>
<td><em>there</em></td>
<td><em>went</em></td>
<td><em>little</em></td>
<td><em>no</em></td>
</tr>
<tr>
<td><em>were</em></td>
<td><em>are</em></td>
<td><em>like</em></td>
<td><em>will</em></td>
<td><em>go</em></td>
</tr>
<tr>
<td><em>as</em></td>
<td><em>see</em></td>
<td><em>come</em></td>
<td><em>get</em></td>
<td><em>here</em></td>
</tr>
<tr>
<td><em>some</em></td>
<td><em>dad</em></td>
<td><em>got</em></td>
<td><em>them</em></td>
<td><em>came</em></td>
</tr>
<tr>
<td><em>day</em></td>
<td><em>house</em></td>
<td><em>your</em></td>
<td><em>too</em></td>
<td><em>home</em></td>
</tr>
<tr>
<td><em>time</em></td>
<td><em>could</em></td>
<td><em>did</em></td>
<td><em>put</em></td>
<td><em>Mrs</em></td>
</tr>
<tr>
<td><em>away</em></td>
<td><em>over</em></td>
<td><em>from</em></td>
<td><em>off</em></td>
<td><em>want</em></td>
</tr>
<tr>
<td><em>bear</em></td>
<td><em>going</em></td>
<td><em>where</em></td>
<td><em>mother</em></td>
<td><em>I'm</em></td>
</tr>
<tr>
<td><em>an</em></td>
<td><em>took</em></td>
<td><em>made</em></td>
<td><em>how</em></td>
<td><em>didn't</em></td>
</tr>
<tr>
<td><em>tree</em></td>
<td><em>who</em></td>
<td><em>monster</em></td>
<td><em>well</em></td>
<td><em>play</em></td>
</tr>
<tr>
<td><em>again</em></td>
<td><em>or</em></td>
<td><em>think</em></td>
<td><em>wanted</em></td>
<td><em>I'll</em></td>
</tr>
<tr>
<td><em>gave</em></td>
<td><em>long</em></td>
<td><em>water</em></td>
<td><em>take</em></td>
<td><em>everyone</em></td>
</tr>
<tr>
<td><em>mouse</em></td>
<td><em>next</em></td>
<td><em>school</em></td>
<td><em>way</em></td>
<td><em>stop</em></td>
</tr>
<tr>
<td><em>must</em></td>
<td><em>box</em></td>
<td><em>new</em></td>
<td><em>been</em></td>
<td><em>find</em></td>
</tr>
<tr>
<td><em>began</em></td>
<td><em>magic</em></td>
<td><em>why</em></td>
<td><em>other</em></td>
<td><em>something</em></td>
</tr>
<tr>
<td><em>say</em></td>
<td><em>found</em></td>
<td><em>tell</em></td>
<td><em>let</em></td>
<td><em>through</em></td>
</tr>
<tr>
<td><em>told</em></td>
<td><em>he's</em></td>
<td><em>shouted</em></td>
<td><em>there's</em></td>
<td><em>soon</em></td>
</tr>
</tbody>
</table>

Copyright © 2007 MULTILIT Pty Ltd. All rights Reserved
The WARL: Grade-based norms

<table>
<thead>
<tr>
<th>Grade</th>
<th>Beginning Yr 1</th>
<th>Middle Yr 1</th>
<th>Beginning Yr 2</th>
<th>Middle Year 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Score</td>
<td>35</td>
<td>55</td>
<td>60</td>
<td>72</td>
</tr>
<tr>
<td>Bottom 25th %ile cutoff</td>
<td>18</td>
<td>39</td>
<td>46</td>
<td>58</td>
</tr>
</tbody>
</table>
The WARL: Reliability & Validity

Alternate Forms Reliability: 0.80-0.97

Criterion Validity:

• Martin & Pratt: 0.66
• BURT: 0.79
• TOWRE:
  - Sight words: 0.95
  - Phonemic Decoding: 0.79
• WARP: 0.85
The WARL – Progress monitoring of individual students
**Summary of curriculum-based measures for use within a RtI model**

<table>
<thead>
<tr>
<th>Year</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1</td>
<td>WARL</td>
</tr>
<tr>
<td>Year 2</td>
<td>WARL/WARP</td>
</tr>
<tr>
<td>Year 3</td>
<td>WARP</td>
</tr>
<tr>
<td>Year 4</td>
<td>WARP</td>
</tr>
<tr>
<td>Year 5</td>
<td>WARP</td>
</tr>
</tbody>
</table>
Questions?
alison.madelaine@mq.edu.au