The Science and Practice of Early Years Evaluation
Saskatoon, Saskatchewan
June 18, 2010
J. Douglas Willms
University of New Brunswick
Skill- Building = Success
Children acquire **skill-based assets** as they mature, and these assets are the **building blocks** for future successful development.

We need to change our focus from demographic risk factors to the outcomes that are critical for successful child development.
<table>
<thead>
<tr>
<th>Birth</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Language Development</td>
<td>Readiness Skills</td>
<td>Learning to Read</td>
<td>Reading to Learn</td>
<td>Engagement</td>
<td>Aspirations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Language Development</td>
<td>Readiness Skills</td>
<td>Learning to Read</td>
<td>Reading to Learn</td>
<td>Engagement Healthy Lifestyle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Language and Communication**

**Social and Emotional Development**

- Stimulating environments
- Exposure to a quality and quantity of language
- Opportunities to socialize with other children
Children make better progress in school if they enter kindergarten with strong language and communication skills

Language development depends on an exposure to a quality and quantity of language
Awareness of Self and the Environment (General Knowledge)
Social Skills, Behaviour, and Approaches to Learning
Cognitive Skills
Language and Communication
Physical Development

Parenting Styles
Family Functioning
Maternal Depression
Engagement
Research supports the crucial roles of cognitive and language skills for learning to read:

- letter identification \( (r = 0.53) \)
- concepts of print \( (r = 0.49) \)
- picture naming \( (r = 0.49) \)
- sentence recall \( (r = 0.49) \)
- phonological awareness \( (r = 0.42) \)
- receptive language \( (r = 0.38) \)
- expressive language \( (r = 0.37) \)

(review by Scarborough, 1998)
Reliability, Validity, and Bias
(measuring what counts)
Reliability = consistency of measurement

- Across time
- Among testers
- Across settings
The Science

- Pre-post correlation and its proxies
- Factors contributing to reliability
- Factors contributing to unreliability
Tying Shoelaces

Prevalence of Task Completion (%) against Child's Age
Tying Shoelaces

Prevalence of Task Completion (%)

Child's Age

0 10 20 30 40 50 60 70 80 90 100

2 3 4 5 6 7 8 9 10
Tying Shoelaces

Prevalence of Task Completion (%)

Child's Age

\( d = 0.5 \)

\( d = 3.0 \)
Test Information and Score Distribution
Challenges

• Assessment time
• Coverage
• Training
• Settings
The Practice

- Pilot tests
- Replacement of items
- Large-sample results
Validity = meaning

• Meaning of the construct
• How it is interpreted
• Actions taken
• Consequences of those Actions

A test is not valid or invalid; validity depends on how scores are interpreted and used, and the consequences of those decisions.
• Expert review
• Correlations with known instruments
• Correlations with known groups (e.g., age)
Challenges

• Links between interpretation and actions, and actions and consequences
• Multiple dimensions within measures
• Mistaken interpretations
• Changing practices
• New applications
The Practice

- Factor analysis
- Focus groups on interpretation
- Large-scale demonstrations
**Bias** = inaccurate comparisons of cultural groups

- Inconsistent difficulty of test items
- Differing responses to testing methods
- Inconsistent meaning of the construct
- Inappropriate interpretation
- Irrelevant actions taken
The Science

- Differential Item Functioning
- Small-scale studies
- Focus groups on interpretation
DIF analysis

DIF size

Differential Item Functioning (DIF) analysis shows the comparison between the probability of a correct response for the Focal Group and the Reference Group. The graphs illustrate the difference in probability as a function of the trait.

On the left, the probability of a correct response is plotted against the trait for both groups. The red line represents the Focal Group, and the blue line represents the Reference Group.

On the right, the DIF size is shown. The graph indicates the size of the difference in probability favoring the Focal Group and the Reference Group across different trait levels.

- Difference favoring Focal Group
- Difference favoring Reference Group
DIF size Domain A

Difference favoring aboriginals

Difference favoring anglophones

Awareness of Self and Environment
DIF size Domain D Gross Motor Development

Difference favoring aboriginals

Difference favoring anglophones

Average DIF Size
Challenges

• Over-studied groups
• Historical experiences
• Fears of wrongful use
The Practice

• DIF studies leading to item revision
• Very little attention to meaning and interpretation bias
Concerns about cultural bias usually have less to do with the psychometric properties of tests than they do with how test scores are used and interpreted.
Looking Forward
Early Years Evaluation

Intervention

Inclusion
**Learning** is a function of:

- Quality Instruction
- Appropriate Level of Instruction
- Engagement
- Time
- Context
Don’t Wait to Fail

- Literacy trajectories are established early.
- Children on a low trajectory tend to stay on that trajectory and fall further and further behind.
- The later children are identified as needing support, the more difficult it is to catch up.
The Response to Intervention (RTI) model calls for a tiered approach to intervention that comprises the continuous assessment of student progress in the regular classroom setting, and intervening with students who have learning difficulties or behavioural challenges.
The first tier involves universal strategies for improving student learning with extra support for those experiencing difficulties.

The second tier involves targeted interventions for those who do not respond to tier 1 strategies. The interventions would typically include further, more detailed assessment, a different strategy for delivering instruction, and possibly a modified curriculum.

The third tier aims to address the needs of students that do not respond to tier 2 interventions. At this level, students would normally receive a more intensive individualized program (Canter, Klotz, & Cowan, 2008).
Use Leading Indicators alongside Trailing Indicators
Establish a district PD framework to create the ‘special forces’ team of teachers skilled in improving literacy.
Develop a strategy for increasing learning time with:

- after-school programs
- making the most of June
- summer learning
Inclusive schools are schools in which all students can succeed
All means learners across the categorical boundaries of disability, social class, gender, ethnicity, national origin, sexual orientation, and religion.

*Succeed* means succeeding in terms of *learning*, as well as in terms of *physical, social, emotional, and spiritual outcomes*.

*Inclusion in the school setting* is best achieved when all children are proceeding through the school system with their peers in heterogeneous groups.
What Governments Can Do

**Legislation:** Enact legislation that support children’s rights

**Money:** Provide the supports necessary for inclusion to work

**Information:** Help create an understanding of the importance of inclusive practices

**Support Innovation:** Support schools with inclusive practices
What School Leaders Can Do

Create a New Framework of Understanding:
  At every turn talk about the benefits of inclusion; create and sustain networks that support inclusion

Hire and socialize new teachers: Hire teachers that are committed to inclusion

Alter key structural features of school: Take steps to desegregate the system at all levels
What Teachers Can Do

Embrace and celebrate diversity: Create a new framework of understanding among students and parents.

Become skilled at culturally responsive teaching: Honour different ways of knowing; maintain high expectations for all; develop relationships that are responsive and caring.

Use new forms of assessment: Assess learning (growth) rather than status, and acknowledge a wide range of outcomes.
Every society establishes its own ‘intolerable equilibrium’ for inequalities derived from its social, historical, political and economic context.

However, these can be altered through policy, practice, and reform.