Critical issues in aligning psychometrics with current reading models

Framework for Diagnostic Assessment of Reading Component Skills in Struggling Readers

John Sabatini, ETS

October 2004
Overview

I. Claims
   A. Misconceptions in Reading Research
   B. Projects and populations - Struggling readers

II. Component Reading Models/Theory/Types of Evidence
   A. Types of Tasks
   B. Efficiency of processing

III. Challenges in moving forward - Aligning measures with models and instruction.

IV. A brief glimpse at the challenges of reading comprehension and the struggling reader
Claims

- We have substantive reading theories and agendas (yet to be fully exploited)
  - Models, tasks, evidence, and interventions
- Reading is not algebra, physics or chemistry
- The act of reading is not by its nature a problem solving activity – forced fit
- Efficiency in processing is a hallmark/ indicator of expertise in reading and reading components
- Many learners are robust to instruction
  - Little LSA heads
  - Many teachers are good instructors
- Validity still rules in diagnostic models, but the rules are different about evidence – the mantra
- Suggests/mandates different assessment design, delivery, and analyses – GROWTH MODELING
The perceptual span grows wider as skill improves as eyes glide continuously across the page.

Dyslexia is primarily a visual perceptual deficit.

Dyslexia is an ‘unexpected’ discrepancy between IQ and reading achievement.

Word recognition is a psycholinguistic guessing game.
Eye-movements of Skilled Readers (Kieras & Just, 1984; Rayner, 1997)

- Eyes move in jumps - fixation, saccade
  - no information is processed during saccade
- Perceptual span is about 8 letter spaces
  - every 1.1 words is fixated (10 to 15% regression)
  - skilled readers spend less time per fixation than less skilled
- Average fixation is about 225 ms (4 wrds/sec)
  - knowledge of multi-letter, multi-sound patterns and correspondences
- Word features predict fixation duration
  - eye movements have an internal logic based mostly on words, less on the construction of meaning
Phonemic Awareness tasks can predict pre-K children at-risk of reading difficulty 2-4 years out. Early intervention with Phonemic Awareness training (along with decoding) can prevent or remediate such at-risk children before they reach clinical definitions of reading failure. This is more than descriptive / correlational; it’s explanatory, causal.

Next up – processes underlying reading comprehension.
• Low literate adults – 5-year experimental comparison of word reading interventions with cog and repeated measure battery.
• Adolescents- same as above (MID & HS) with pre-post fMRI [Scarborough]
• Adolescents – groups of comprehension only deficit (RCD) vs. word reading deficit (WRD) (cog battery with fMRI) (Cutting)
• Both - Diagnostic measures for Struggling readers using computation linguistic analyses to identify text features that make text difficult, then diagnostic measures [Deane, Sheehan, Mislevy]
• Accessible Reading assessments ‘for students who have disabilities that affect reading’ (e.g., LD)[Cahalan/Cook]
• Low literate adults in developing countries. [Kirsch]
• Structural model of Efficiency components in EL learners Spanish, Korean, Chinese college freshman (Sawaki)
• Research focus - bottom half of the reading distribution
  – adults first and children
• NOT a Linguist, not an English major, not a psychometrician
  – Psychology trained (behavioral, neuropsych, cognitive)
• NOT achievement/outcome testing
  – Classroom-based/formative
  – assessment/learning
• Psychological Component Measures
  – Published one-on-one (e.g., W-J, TOWRE)
  – Experimental (e.g., response rate, voice-onset)
• Eye movements
• Parallel Distributed Processing/connection simulation modeling
• Psycholinguistic studies
  – LSA and Computation Analyses of text and speech; experimental manipulations
• Neuroimaging (fMRI)
• Educational intervention studies
• Simple View of Reading
  – Hoover & Tunmer (1993): "the simple view makes two claims: first, that reading consists of word recognition and linguistic comprehension; and second, that while each of these components is necessary for reading, neither being sufficient in itself." (p.3)

  – Strucker, Yamamoto, & Kirsch (2003) use a similar component framework when they describe *print components* (e.g., decoding accuracy and fluency) and *meaning components* (e.g., oral vocabulary).
Reading can be thought of as comprised of two component processes: decoding and language comprehension. Neither alone is sufficient.
Guiding principles

• Reading comprehension is built upon a foundation of component skills and knowledge of how one’s writing system works. (Perfetti, 2003)
  – Your language; your script; your social context

• The level of efficiency, automaticity, and integration of component skills is indicative of level and learning potential in reading development
  – Component efficiency is typically indexed by assessing speed/latency or rate of processing, as well as accuracy

• In follow-up studies of children with reading disabilities, it is not uncommon for age-appropriate accuracy to be demonstrated in adolescence or adulthood, yet when speeded tasks are given persisting weaknesses in the efficiency with which their accuracy levels can be attained
Carver, 1997
The Many Strands that are Woven into Skilled Reading
(Scarborough, 2001)

LANGUAGE COMPREHENSION

BACKGROUND KNOWLEDGE
(facts, concepts, etc.)

VOCABULARY
(breadth, precision, links, etc.)

LANGUAGE STRUCTURES
(syntax, semantics, etc.)

VERBAL REASONING
(inference, metaphor, etc.)

LITERACY KNOWLEDGE
(print concepts, genres, etc.)

WORD RECOGNITION

PHONOLOGICAL AWARENESS
(syllables, phonemes, etc.)

DECODING (alphabetic principle, spelling-sound correspondences)

SIGHT RECOGNITION
(of familiar words)

SKILLED READING:
Fluent execution and coordination of word recognition and text comprehension.
Reading can be thought of as comprised of two component processes: decoding and language comprehension. Neither alone is sufficient.
Components of Reading Instruction

INSTRUCTION

Fluency

Comprehension
• vocabulary
• cognitive strategies

Alphabetic
• phonemic awareness
• phonics
Typical Assessment Battery

WRAT
- Reading (Word naming)
- Spelling

TOWRE
- Speeded Word Naming
- Speeded Decoding

CTOPP
- Phonemic Awareness – (Elision and Blending)
- Rapid symbol naming (letters and digits)
- Phonological Memory (nonword repetition and memory for digits)
Typical Assessment Battery

WJIII – Woodcock Johnson III
- Word Identification (Word naming)
- Word Attack (Decoding)
- **Passage Comprehension**
- Reading Fluency
- Oral Language Cluster (Oral comp; Understanding Directions; Story Recall)
- Spelling of Sounds (Phonemic Awareness)
- Picture Vocabulary

**Boston Naming** – Picture Vocabulary
Diagnostic model

Word Identification -> Word Meaning -> Sentence Processing

Sentence Processing -> Reading Comprehension

Reading Comprehension -> Literary Interpretation

Text/ Discourse/ Genre Modeling
Student Model variables

Task/Evidence Model Variables

NOTE: Model Upside Down!

1. Task Mod 1
2. Task Mod 2
3. Task Mod 3
4. Task Mod 4

Word ID
Sentence Processing
Word Meaning
Text/Discourse Processing
Diagnosis Aims - Change

- Identify struggling readers early (at-risk) to prevent or remediate
- Provide information that informs teachers about instruction /interventions (and their own teaching)
- Just in time diagnosis [if instruction is working, less need to diagnose; monitor]
- More severe the case, more profiling [A & B &…].
- Longitudinal profile should improve reliability, validity, and reduce time in subsequent testing.
- Thus, growth models [sensitive to response to intervention]
- Normative developmental models of growth expectations by component.
Challenges

External Validity (Morris, p. 83)
It includes the identification of group differences in terms of:
- external descriptors,
- developmental outcomes,
- treatment effectiveness, or
- etiological bases for the disorders.

Without such differences, the most standardized and reliable classification system is of no clinical, empirical, or predictive use.
Challenges

Francis et al (1994)

The concept of change is fundamental to any comprehensive definition of learning disability. To appreciate this assertion, consider the role of change in the validity of any inference regarding the presence or absence of learning. Because learning is, by definition, the acquisition of skill or knowledge, it is difficult to demonstrate that an individual has learned unless it is possible to demonstrate that the behavior has “changed”.
Francis et al (1994)
Thus, the concept of learning disability must imply some departure from normality in the process of change. A difference may exist in
• the rate at which new material is acquired or
• the duration of acquisition period,
• what material is acquired, or
• how material is acquired.
Regardless of the specific difference involved, the inclusion of the term learning in the label *learning disabled* has implications for change over time.
Challenges

Francis et al (1994)
In contrast, current approaches to the definition and diagnosis of learning problems typically focus on differences in outcomes at specific points in time and do not directly examine change.
The valid measurement of behavioral change over time... is virtually impossible at present because few instruments have the necessary scaling properties to satisfy conditions for longitudinal measurement (Francis et al., in press from Lyon & Moats, pg 6)
“A major similarity between mental retardation and learning disabilities is that their definitions at this juncture are inherently psychometric.

More specifically, definitions based on

- etiology,
- responsiveness to treatment, or
- biological correlates presently are not possible.” (Fletcher et al, 1993, p.33)
Diagnostic measurement subsumes summative assessment. How?

Thursday: Cognitive Diagnosis built upon:
• Small set of items – mostly multiple choice
• One test session
• Heavy investment in quality of each item
• Integrated (thinking items) that require extracting skills
• Except tutoring/adaptive models.
Imagine - Cognitive Classroom Diagnosis built upon:

- Predictive variables (e.g., pho awareness)
- Multiple, continuous classroom data collection (computer) [So often, it’s not worth gaming.]
- Longitudinal, growth profile [CBM]
- Response to intervention/instruction
- Voluminous data sets of small tasks from which we can sample.
- Variety of task and evidence types – eye movement; speech recognition, response rates, iterative loops of items [Reading process]

Then demonstrate that we can predict summative measures from modeling formative growth. [Validity]
Summary

There is substantive theory emerging from an interdisciplinary convergence of research sources. This includes testing theory by triangulation, training, & instructional intervention in experimental settings with diverse populations.

The measures, methods and tools used to build and contrast theoretical constructs are a rich source for assessment designs.

School-based diagnostic, formative, adaptive, longitudinal measurement of struggling readers would benefit from more sophisticated growth models fit against developmental empirical data. [BMI]

With such an assessment system in place, so-called formative/diagnostic measurement could subsume summative tests.