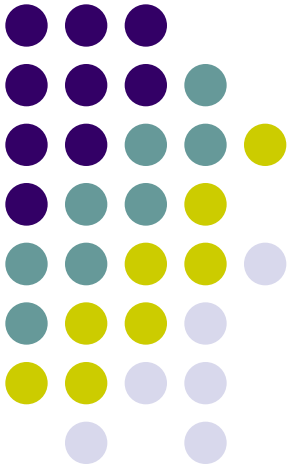
Diagnostic Assessment

Katherine Stahl
kay.stahl@nyu.edu
January 13, 2014

NYS RTI TAC Webinar Series



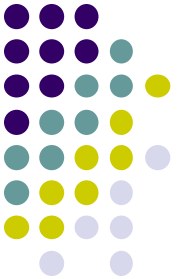
To access presentation:

http://www.nysrti.org

How much is enough?

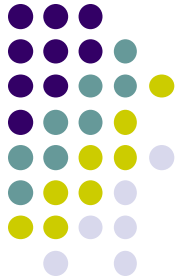
- How much is too much?
- What types of assessments should we use?
- Can we use CBMs to inform instruction if we record errors and perform a formal miscue analysis on those errors?
- Can we use running records as screening instruments and for progress monitoring if we time the reading?
- How do we use diagnostic assessment to increase student performance on the new CCSS ELA tests?

What do we want our assessments to do?



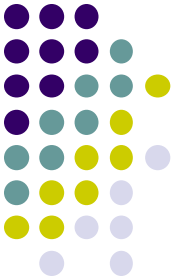
- Measure the achievement of students
- Determine student needs
- Trace the student learning trajectory
- Evaluate the effectiveness of instruction

What is the role of assessment in an RTI model?



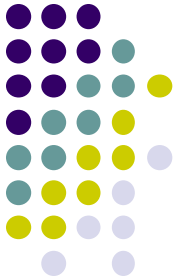
- A move away from a single snapshot determination to a GROUP PROCESS determination.
- Screening measures are one “indicator” of children who may need additional support in literacy.
- Historically, students had to “fail” for a sustained period before eligible for “special” services.
- Progress monitoring in ALL settings holds the promise of growth for ALL students regardless of context.

Evaluating assessments in an RTI model



- Psychological tests and CBMs are not designed to inform instruction.
- Often our lowest performing children are operating at frustration levels on CBMs, so it is not a true picture of function.
- While one-minute measures insure the sensitivity to measure change in a short period of time, they are not specific enough to generate prescriptions for instruction.

Categories of Assessments



Types

- Summative
- Formative

Functions

- Screening
- Diagnostic
- Progress Monitoring
- Evaluative

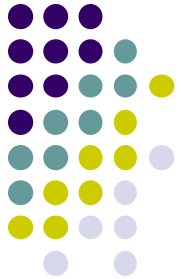
Summative Assessments

- Snapshot of a student at a particular point in time to determine that student's performance relative to a standard
- Occur after instruction
- Used to evaluate the effectiveness of programs, achievement of school goals, curriculum alignment, instructional techniques, student placement in programs

Examples of Summative Assessments

- NYS ELA
- F&P Benchmark (Text Level)-holistic
- District benchmarks
- Unit or chapter tests
- WTW/Ganske Spelling Inventory Power Score
- Woodcock-Johnson
- Peabody Picture Vocabulary Test

Formative Assessments

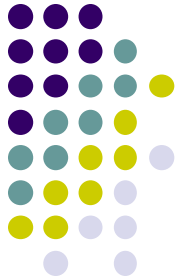


- Provide the information needed to monitor and adjust instruction
- May occur in a variety of formats
- Administration and scoring occur along a continuum of standardization of procedures
- Should inform both teachers and students about ongoing movement toward learning goals

Formative Assessment

- Many formative assessments have high positive consequential validity.
- What is the consequence of assessment process?
- Act of assessment yields positive consequences
  - Instructional
  - Learning

Formative Assessments

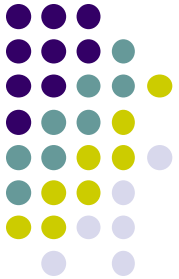


- May be difficult to separate from instruction because it is used to inform instructional process
- Often teacher constructed
- Students are involved in cyclical process
- Usually diagnostic

Examples of Formative Assessment

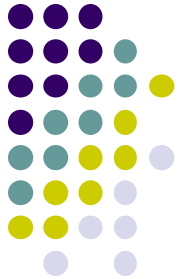
- WTW qualitative stage assessment
- Observations/anecdotal records
- Questioning strategies
- Self and peer assessment
- Daily running records and MSV analysis

Some Specific Functions of Assessment



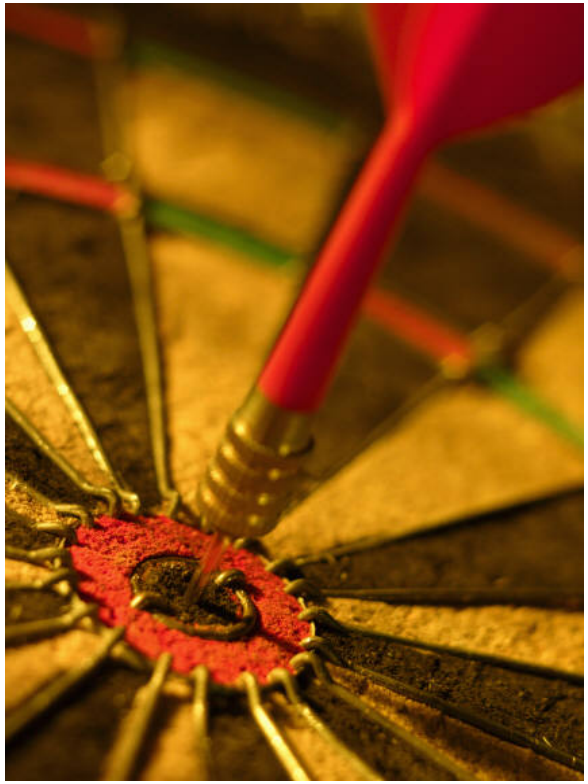
- Screening
  - Given at beginning of the year to quickly identify children who may need additional help
- Diagnostic
  - An in-depth follow-up to screening
  - Results inform instruction
- Progress monitoring
  - Short measures given throughout the year to make sure all children are developing at an adequate rate
- Evaluation-Outcome
  - Given at the end of the year to make sure that children made adequate progress

Developing a Tiered Model of Assessment Roles



	Tier 1	Tier 2	Tier 3 and Sp. Ed.
Screening			
Instructional Diagnostics			
Progress Monitoring			
Diagnostic 2		***	***
Other			

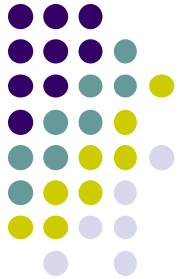
Screening measures



Common measures

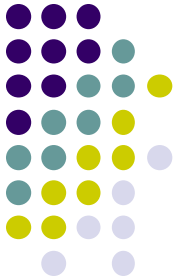
- NYS ELA
- Achievement tests- SAT9, Terra-Nova
- CBM: DIBELS/AimsWeb
- Mini IRI (grade level passage) or Benchmark Kit (F&P, DRA)

Diagnostic Measures



- Extended IRI process
- Developmentally specific tools within each NRP pillar
  - Phonics measures
  - Developmental spelling inventories
  - Think-aloud tasks
  - Retellings, summaries, a range of reading response formats with rubric evaluation
  - Strategy indices
  - Selective vocabulary assessments

Progress Monitoring



- A means of microscopically examining student achievement, tracing individual learning trajectories and evaluating the effectiveness of instruction within a fairly short duration of time
- Focus = achievement level and rate of progress

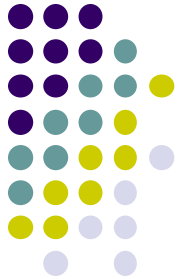
CBMs

- Standardized
- Constant, sensitive measure
- Each weekly test is equivalent
- General outcome = Span the school year
- Fixed time
- DIBELS, AIMSweb

Three Goals of Reading Instruction

- Comprehension and Learning from text
- Automatic word recognition
- Motivation and Appreciation

The primary purpose of reading is to comprehend.



- Automatic word recognition
- Language comprehension
- Intentional strategies used in flexible ways specific to purposes for reading and text

The Cognitive Model

Phonological Awareness

Print Concepts

Decoding and Sight Word Knowledge

Fluency in Context

Automatic Word Recognition

Vocabulary Knowledge

Background Knowledge

Knowledge of Text and Sentence Structures

Language Comprehension

Reading Comprehension

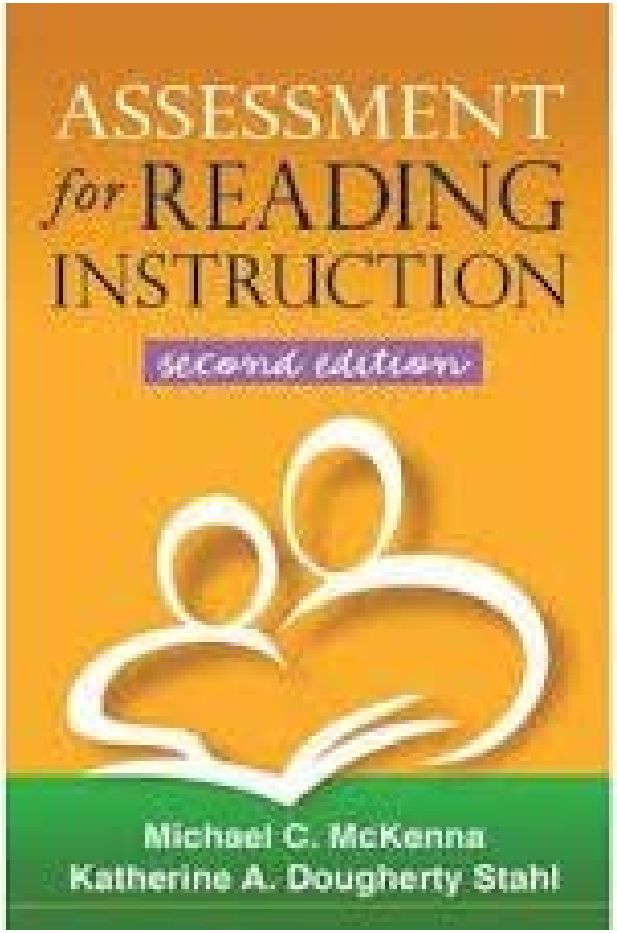
Strategic Knowledge

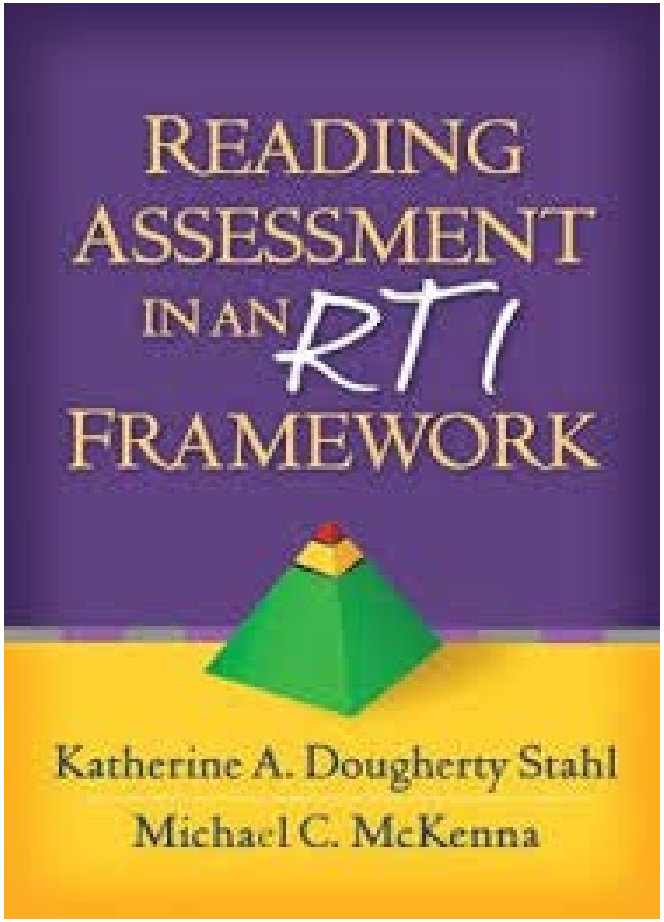
General Purposes for Reading

Specific Purposes for Reading

Knowledge of Strategies for Reading

Two Books that I Will Reference





Automatic Word Recognition

- Phonological awareness
- Decoding
- Sight word knowledge
- Fluency

Phonemic Awareness

SCREENING

CBM Initial Sound

CBM Segmentation

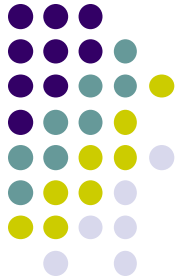
DIAGNOSTIC

Tasks based on NRP

(McKenna & Stahl, p. 98)

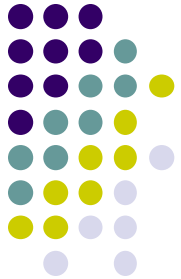
Invented Spelling
(Dictation Task)

Phonological Awareness



- Caution with older students
- Invented spelling can provide insights
- NRP recommends that instruction range from 5-18 hours, certainly no more than a total of 20 hours (6.5 minutes/day)

Decoding



- Become familiar with the stages of decoding and spelling development.
- Use caution-do not overemphasize decoding instruction with older students.
- Differentiate instruction based on individual developmental assessments of decoding and spelling stages.

Important stages

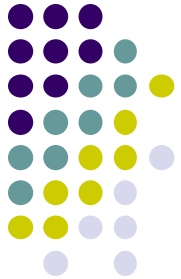
Word recognition

- Emergent
- Phonic
- Analogies
- Syllabic
- Morphemic

Spelling

- Emergent
- Letter name
- Within Word
- Syllable & affixes
- Derivational relations

Differentiating Instruction



- Remember, that the NRP findings support explicit and systematic phonics instruction in kindergarten and first grade.
- Children should be taught material at their developmental level-not one size fits all.
- Only diagnostic assessment tells us what those needs are.
- In an RTI model, children in Tier 3 and Special Education should absolutely NOT be placed in a one-size-fits-all phonics program. They need individual educational support from a highly trained expert teacher driven by diagnostics and weekly PM.

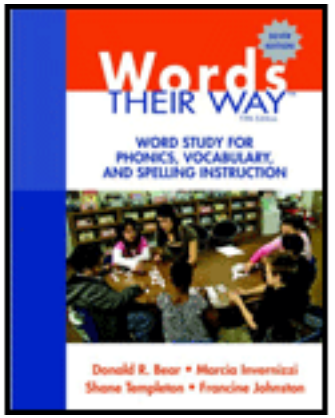
Screening Assessments

- CBM- DIBELS or AIMSweb
- Letter Naming
- Nonsense Words
- TOWRE Phonemic Decoding

Diagnostic Assessments

- Systematic Phonics Inventories (McKenna/Stahl Inventory)
- Z-Test (McKenna/Stahl)
- Miscue analysis of word lists and connected text reading

- Developmental Spelling Inventory

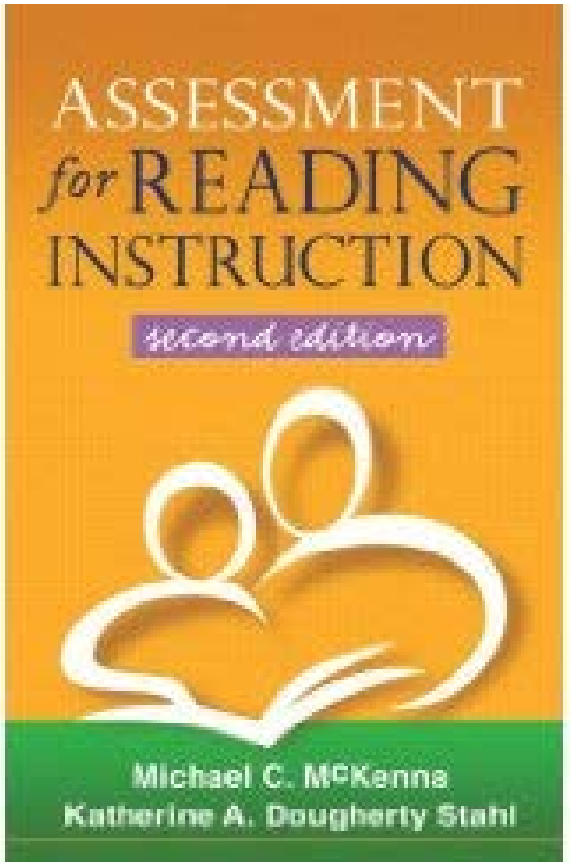


  - *Words their Way*

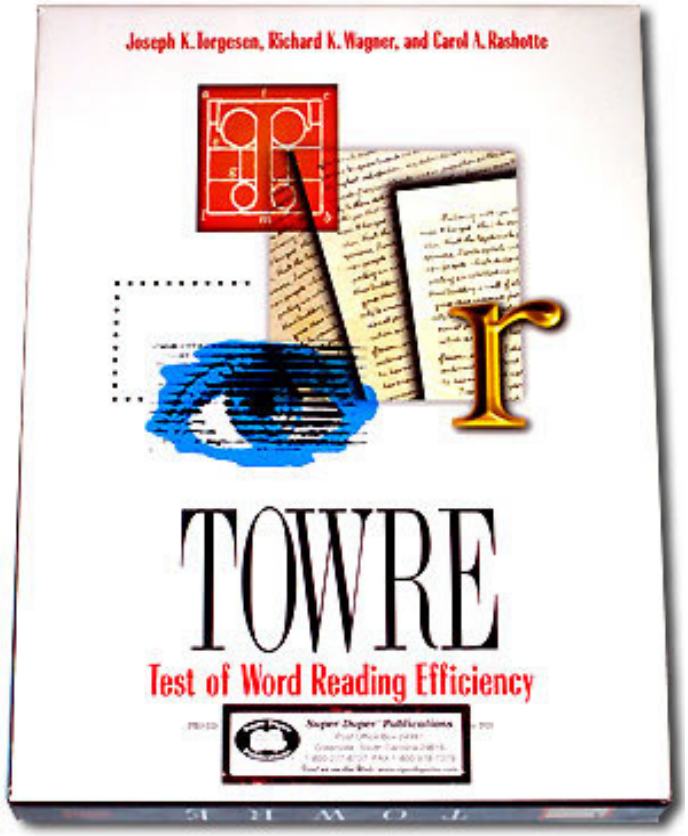
    (Bear, Invernizzi, Templeton, Johnston)
  - *Word Journeys* (Ganske)

Assessments

- Phonological Awareness Tasks
- Phonics Inventory
- Spelling Inventory
- Sight Word Lists
- Multiple Dimensions of Fluency Scale

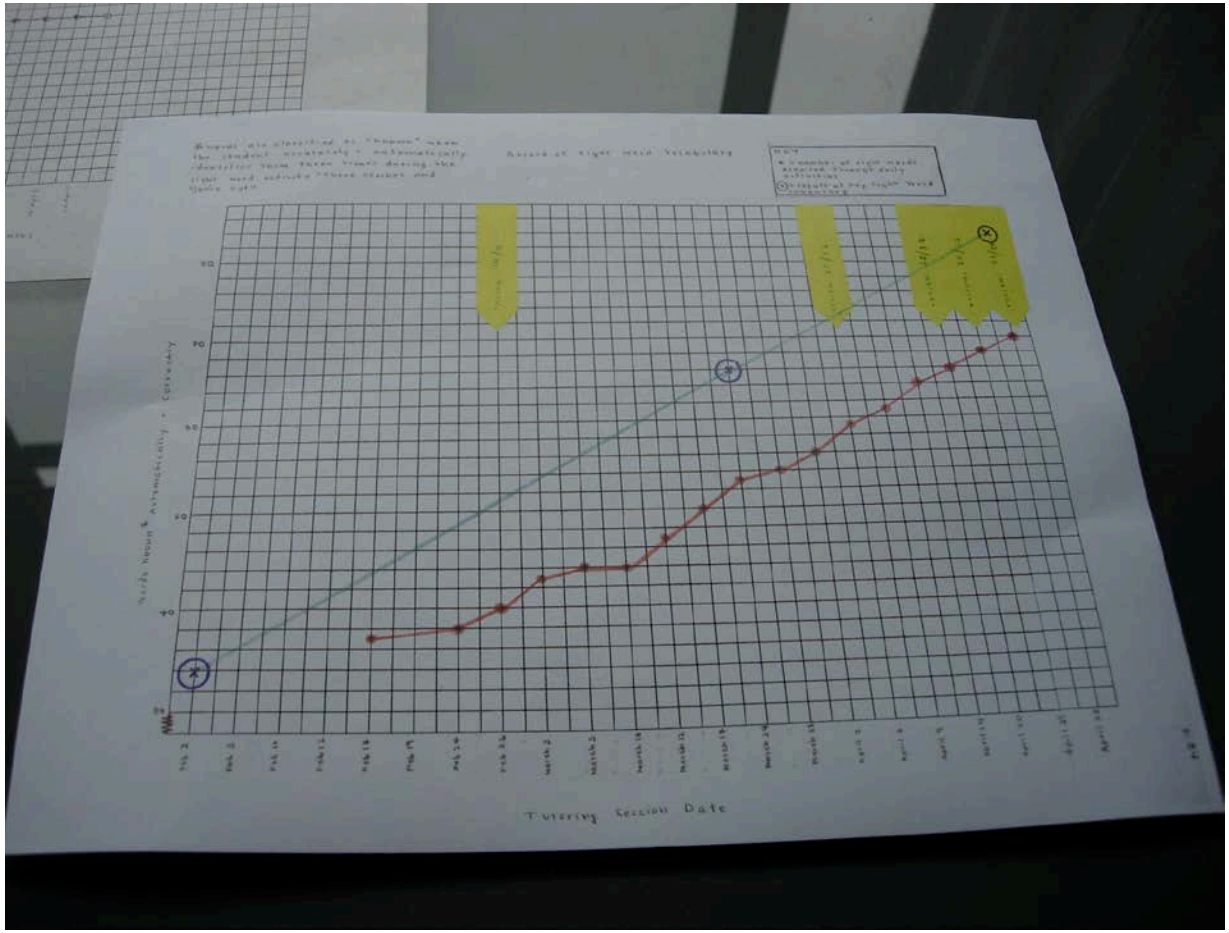


Sight Word Knowledge



- SCREENING
  - TOWRE (>Gr. 3)
- DIAGNOSTIC
  - Fry or Dolch List
- PROGRESS MONITORING
  - Intervention Central
  - Fuchs' Package

High Frequency Words



Fluency

- CBMs (Oral Reading Fluency)-Not retelling/comprehension
- Timed reading on IRI or other samples of connected text --Words Correct/Minute
- Leveled text running record

Fluency Rate Norms

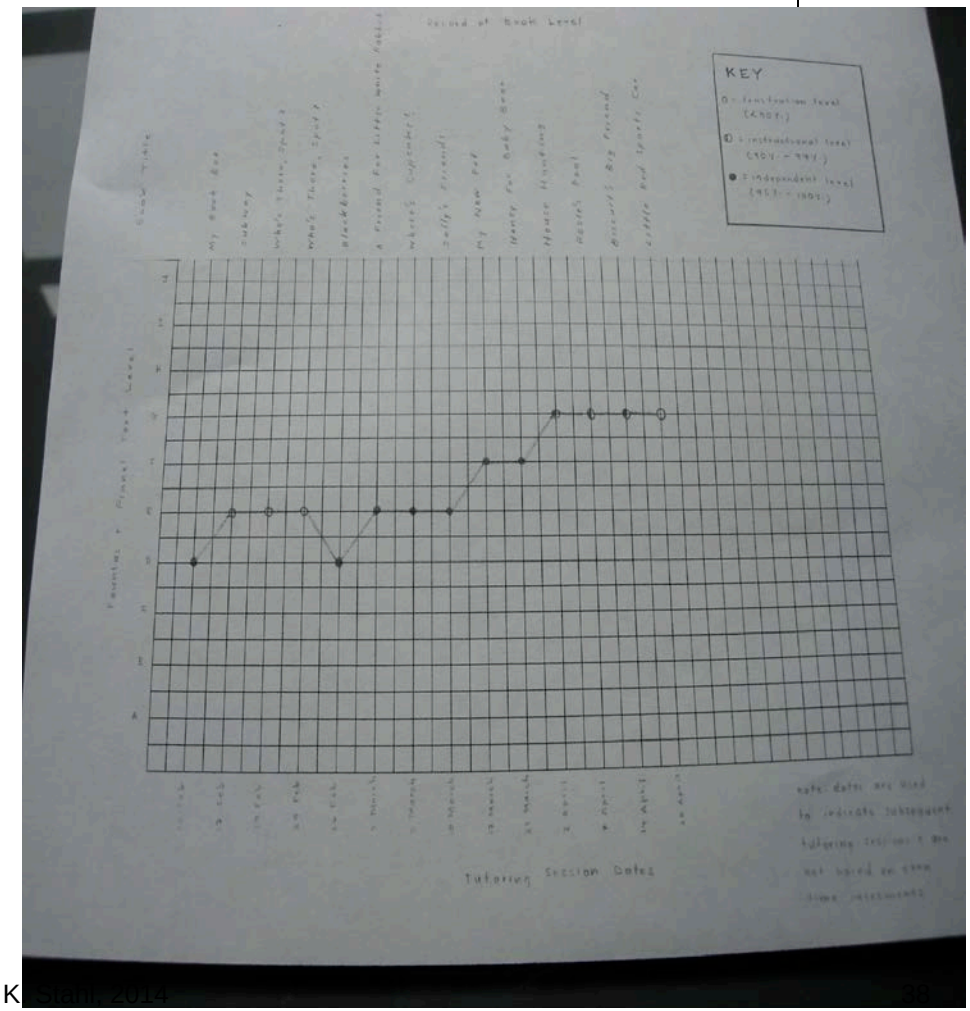
- 2006 Hasbrouck & Tindal Oral Reading Fluency Data Table
- To raise fluency wcpm, differentiate instruction based on TOWRE, word recognition, and sight word diagnostic information-NOT DIBELS/AIMSweb.
- Remember, DIBELS is a screening tool.

Prosody

- Expression, pitch, stress
- Use a prosody rating scale (NAEP or Rasinski & Zutell Multidimensional Fluency Scale) in conjunction with oral reading wcpm score.
- Multidimensional Fluency Scale can be used to inform instruction and is sensitive enough for progress monitoring.

Progress Monitoring

- ORF
- Text Level
- Prosodic Features



Let's Shift Gears

- Constrained abilities are fairly linear and with instruction children develop mastery within a few years.
- Unconstrained abilities are multidimensional, incremental, context dependent and develop across a lifetime.

Paris, S. (2005). Reinterpreting the development of reading skills. Reading Research Quarterly,40, 184-202.

Stahl, K. A. D. (2011). Applying new visions of reading development in today's classrooms. The Reading Teacher, 65, 52-56.

Continuum: Constrained to Unconstrained Abilities

(Paris, 2005)

Phonemic Awareness

Fluency

Vocabulary

Phonics

Comprehension

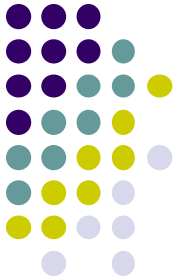
Constrained

Unconstrained

Language Comprehension

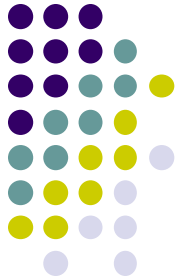
- Background Knowledge
- Meaning Vocabulary
- Knowledge of Structure

National Reading Panel: Vocabulary Findings



- Most studies in NRP were conducted in grades 3-6, followed by Pre-K and K.
- Instruction makes a difference in vocabulary learning.
- Standardized assessments only provide global baseline…view results tentatively.
- Teacher-generated assessments that match instruction are recommended. It is critical to use more than a single measure.

Evaluating Vocabulary Measures
(Pearson et al., 2007; Read, 2000; Stahl & Bravo, 2010)



This continua might be used to design and evaluate vocabulary measures.

- Discrete - embedded
- Selective - comprehensive
- Decontextualized- contextualized

Screening

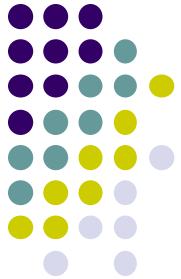
- Apply cautiously- New special education driven subject area tests (Espin et al., Vannest et al.)
- Espin et al. used a collection of vocabulary from specific units-short term, definitions only, maze or matching.
- More recently, work in SS and Science curriculum vocabulary banks. Used long term for PM. Computer generation reliant. (Vannest et al., 2009)
- Matching format. Both definitions and application. Computer administered. No time limit.

Vocabulary and Knowledge

(Stahl & Bravo (2010) Contemporary classroom vocabulary assessment for content areas; Reading Teacher, 63, 566-578)

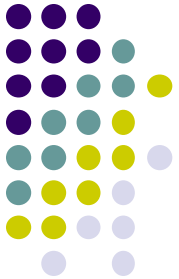
- Vocabulary Knowledge Rating Scales
- Vocabulary Recognition Tasks
- Anecdotal Notes on Conversations
- Comparative analysis of writing samples
- Cognate charts

COMPREHENSION: The Challenge



- NYS ELA IS NOT A DIAGNOSTIC AND SHOULD NOT BE USED FOR GROUPING.
- ITEM ANALYSIS of NYS ELA IS NOT EFFECTIVE USE OF TIME. (Stahl & Schweid, 2013)
- CBMS ARE NOT DIAGNOSTICS AND SHOULD NOT BE USED FOR GROUPING.

Screening Comprehension Measure: IRIs



- Informational and Narrative Passages
- Good reliability and validity ratings
- Decent form to form consistency
- More difficult and time consuming to administer

Derive diagnostics from deep analysis of the IRI components- fluency, retelling, Q/A ability

Written Responses to Reading

- Position taken in response to the prompt question
- Support from texts
- Use state rubric for teaching and testing

Progress Monitoring

- AIMSweb or other commercially produced Maze measures
- Retelling rubrics that consider both idea units and story grammar elements (Be discriminate, this is time-consuming)
- Written response units and NYS rubric

Strategy Interviews

- Burke Reading Interview-Perceptions and purposes of reading
- Index of Reading Awareness (Jacobs & Paris)
- Textbook Interview
- Think-alouds (QRI-Leslie & Caldwell)

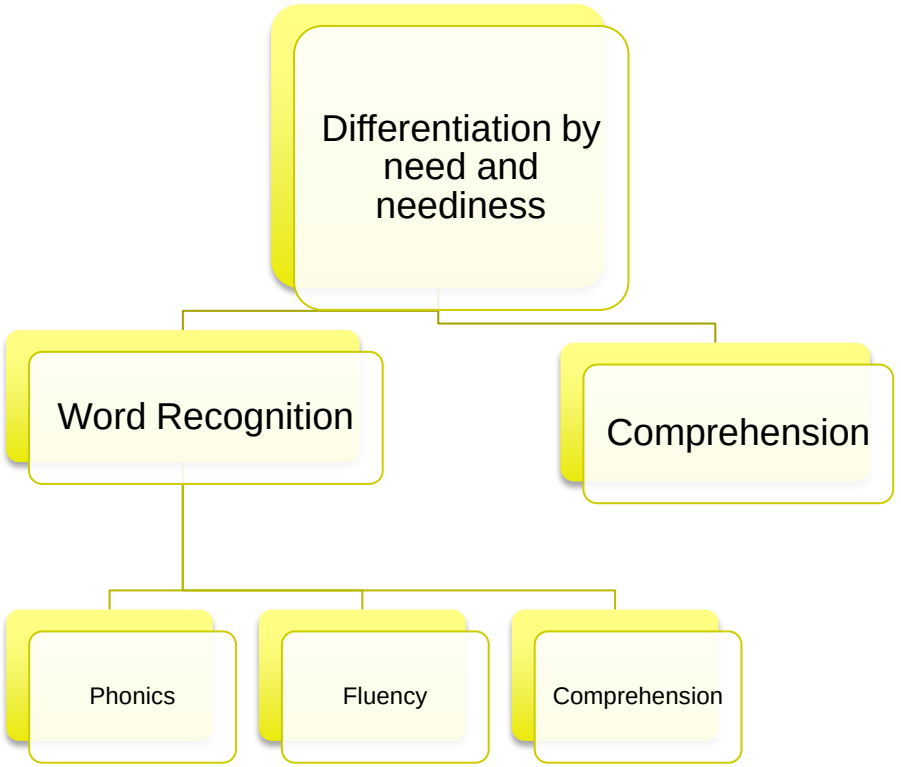
Affective factors

- Journals
- Questionnaires
- Attitude Surveys

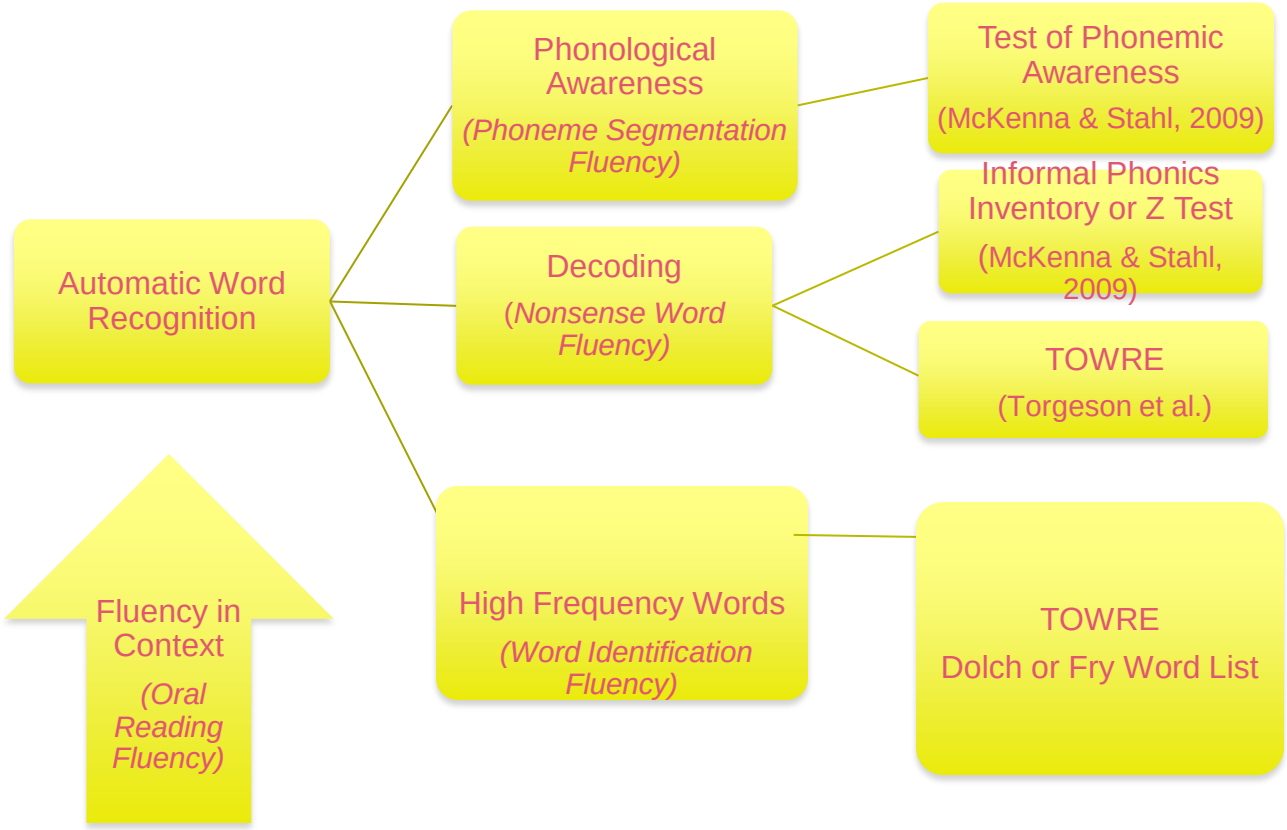
- Uninterested
  - Expand student choices
  - Use Book Series
  - Add a social aspect
  - Model enthusiasm
- Poor Decoding
  - Use assisted reading

Identifying and Grouping for Instruction Beyond Gr. 2

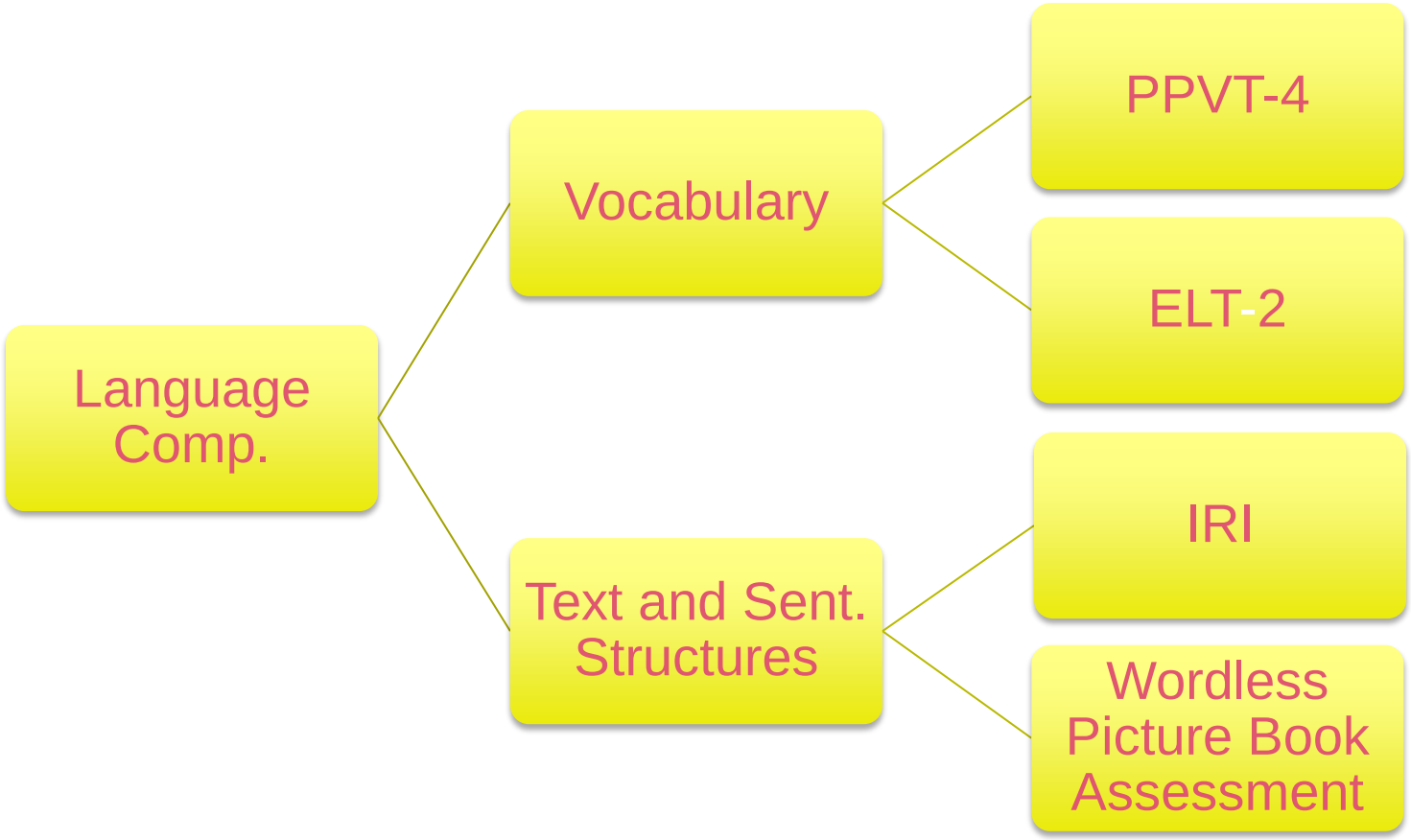
- Use triangulation of ELA, CBM ORF, WTW Elementary Spelling Inventory and IRI (retelling and questions/Lexile correlated) to identify needs and the neediest.



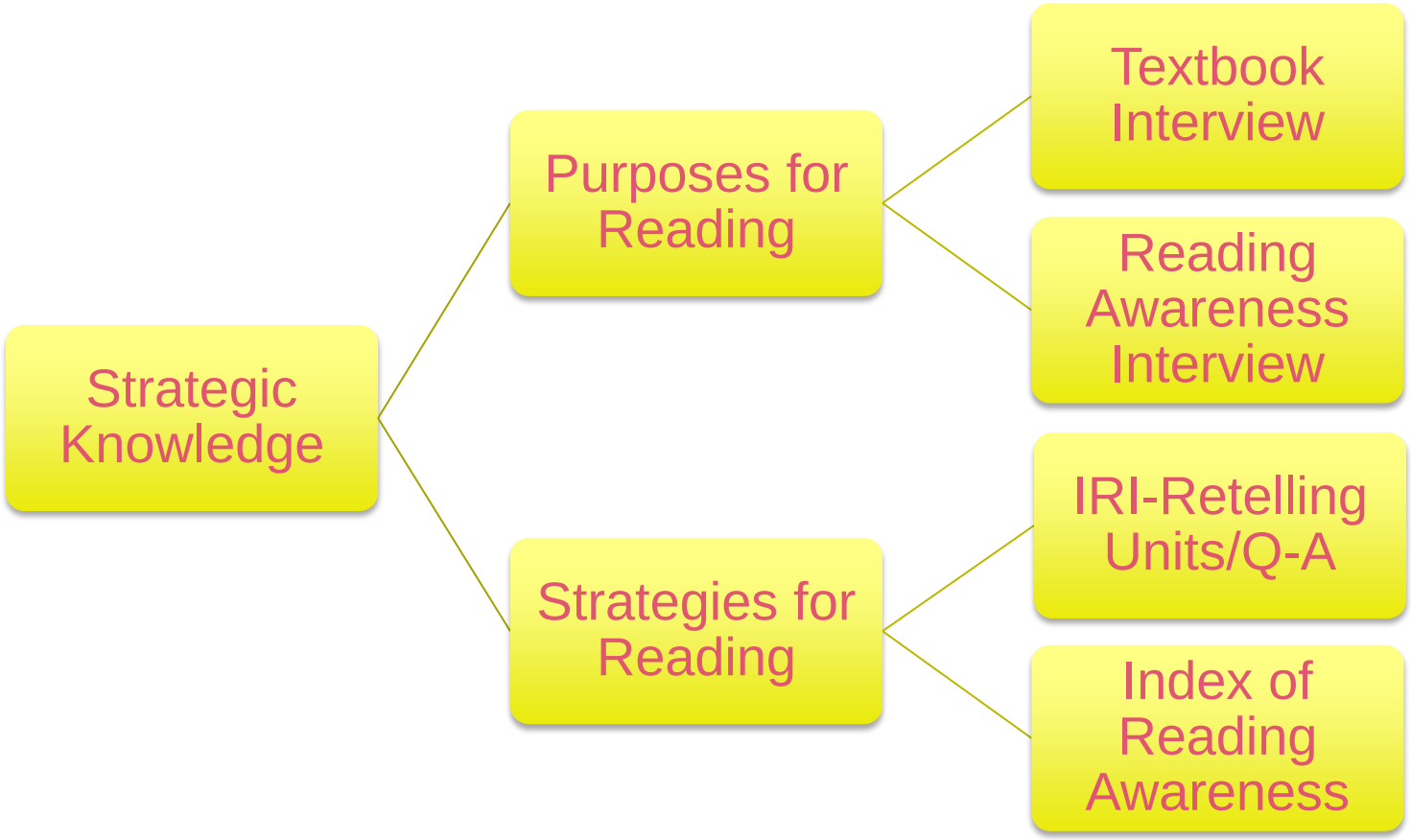
Assessing for Tier 3 (Stahl & McKenna, 2013)



Assessing for Tier 3



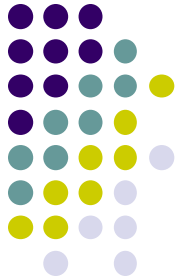
Assessing for Tier 3



Difficult Conversations

- Training
  - Based on school needs assessment, what kind of PD is needed and for whom.
  - All teachers may not need deep knowledge of all assessments. Consider assessment overview PD for all and separate administration PD for few.
  - Since special ed. teachers are teaching children who need individualization, knowledge of "programs" is inadequate. Diagnostic expertise and research-validated techniques are demanded.

Difficult Conversations



- Transparency
  - Planned data sharing for every single child in Tier 2 and Tier 3 with classroom teacher, literacy specialists and special education teachers with others as periodic visitors/viewers.
  - Remember, this is important. Data will be used to determine instructional status.

Fall in  with data

- Use it to help you know your students.
- Use it to help you know yourself.
- Use it to make your life and the lives of your kids easier.