The reading wars through the NRC (2001) looking glass: Reflections on testing for reading and reading for testing

Brent Duckor, Ph.D. & Agnes Hodi, Ph.D.
UC Berkeley BEAR Talk
October 15, 2019
The Reading Instruction Landscape

- “In the beginning” or at least since colonial times a highly synthetic (from letters to words) alphabetic approach prevailed in teaching reading. (Pearson, 2004; Schoenfeld and Pearson, 2009)

- “Phonics” as we now call it was used to teach the correspondence between the sounds and letters that represent them.

- In the 1920s, the appearance of instructional alternatives to reading emerged.

- “Whole-language” as we now call it was conventional wisdom between 1930-1970.

- But whole-language was challenged (Flesch, 1957) during Sputnik era, and then came under serious pressure with the Phonics movement in 1980s.
The “Reading Wars”

“A popular metaphor for describing tensions among different camps in reading curricular history—a struggle between competing goals, a war between opposing ideological forces, or a pendulum that swings back and forth between opposing curricular viewpoints."

Schoenfeld and Pearson, 2009, p. 561 [emphasis added]
Embedded in this struggle over reading instruction and delivery are...

- thematic tensions (bottom up v top down)
- value conflicts (academic vs. vocational purposes)
- philosophies of education (constructivist, behaviorist, cognitive)
- teacher capacity building/”upskilling” the workforce
- industry demands and publishers’ interests in scalable products and services
- achievement indicators at multiple levels of accountability “system”
  - local school site
  - district
  - state
  - national
  - international
# Two Sides in the Reading Wars: “Camps”

<table>
<thead>
<tr>
<th>Phonics</th>
<th>Whole-language</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Reading as a perceptual process</td>
<td>• Reading as a language experience</td>
</tr>
<tr>
<td>• Skill-building hypothesis: conscious learning of rules</td>
<td>• Comprehension-hypothesis: skill mastery will be inferred from / the by-product of the meaning-making process</td>
</tr>
<tr>
<td>• Skill-centered curriculum</td>
<td>• Meaning-based curriculum - integrationist view</td>
</tr>
<tr>
<td>• Heavily reliant on direct instruction</td>
<td>• Skills are better caught in the act of reading than taught directly and explicitly by teachers</td>
</tr>
<tr>
<td>• Well-defined and extensively researched constituents of reading</td>
<td>• Starts out from word form-meaning associations</td>
</tr>
<tr>
<td>• letter-sound associations</td>
<td>• comprehension and interpretation</td>
</tr>
<tr>
<td>• phonological awareness</td>
<td>• Method for observation: qualitative study and ethnographic research</td>
</tr>
<tr>
<td>• Method for observation: quantitative study and psychometrics</td>
<td></td>
</tr>
</tbody>
</table>
“Balanced Literacy” (BL): A happy compromise?

Instead of picking a side/camp, BL advocates say we should be flexible.

BL advocates say we should ask what aspect of either approach—**Phonics and Whole Language**—leads to which outcomes and why? (Chall, 1967)

BL emphasizes discrete “**decoding**” AND “**comprehension**” skills.

BL demonstrates a **behaviorist commitment** to “direct instruction” and repetition AND a **constructivist orientation** toward utilizing “authentic” literature and language experiences.

Nonetheless, some say BL is “a confusing mishmash of conceptions” (Pressley, Roehrig, Bogner, Raphael and Dolezal, 2002; Pressley, 1998).
Reframing the Reading Wars:
Aren’t these the assessment wars by other means?

There are several “fronts” in this struggle for reading approaches.

The battles for instructional strategies have been supplanted by the “theatres” of assessment, accountability and testing since NCLB (2001).

No one wins the reading wars without having a superior technology of assessment, hence, the stalemate and “one off” victories.
The battle for Phonics as part of the “assessment wars”

- NCLB Reading First Initiative: recipients of the grant are required to administer scientifically-based screening and diagnostic assessments to determine which kindergarten through third grade students are at-risk for reading difficulties.

- Schools need research-based, valid and reliable instruments.

- Hard-data is needed in data-driven education systems.

- Phonics produces measurable results.
The battle for Whole Language as part of the “assessment wars”

- Advocates have rejected the conventions of formal standardized testing and assessment. They argue that traditional “paper and pencil” assessments fail to capture students’ motivational development and higher-order thinking competencies (Guthrie, Van Meter, Mitchell, & Reed, 1994, p. 266).

- Possible solutions have included performance assessments and authentic classroom assessments (Garcia and Pearson, 2004).

- These emphasize judgements of process rather than product/outcome.

- The battle cry for “authentic assessment” a term intended to connote the relationship of an assessment task to application in everyday situations and to distinguish this enterprise from "contrived" formal assessments (Garcia and Person, 2004)
We need to carefully examine

Logic of
Reading Assessment

Observation

Cognition

Interpretation
## Instrument Contexts, Purposes & Data Uses

<table>
<thead>
<tr>
<th></th>
<th>PISA</th>
<th>CAASPP/SBA</th>
<th>LEXILE</th>
<th>DIBELS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Context</strong></td>
<td>International</td>
<td>California</td>
<td>National-International</td>
<td>National</td>
</tr>
<tr>
<td><strong>Purpose</strong></td>
<td>workplace readiness to enter labor market</td>
<td>within school “growth”/progress; college and career readiness</td>
<td>within school-level “growth”/progress; college and career readiness</td>
<td>within school “growth”/progress</td>
</tr>
<tr>
<td><strong>Data Uses</strong></td>
<td>inform educational policy</td>
<td>designed to support instruction and give teachers better information about their students’ progress</td>
<td>provide complementary procedures for measuring people and text</td>
<td>instructional decision making and placement</td>
</tr>
<tr>
<td><strong>History</strong></td>
<td>2000</td>
<td>2015</td>
<td>1996</td>
<td>1996</td>
</tr>
</tbody>
</table>
Unpacking each instrument

1. Theory of cognition and Construct definition of reading (skill, proficiency, etc.)

1. Observational strategy and Items design

1. Outcome space and score interpretation including evidence for validity and reliability
What does the NRC mean by cognition?

“A model of how students represent knowledge and develop competence in the subject domain”
(NRC, 2001, p. 2)

“A theory or a set of beliefs about how students represent knowledge and develop competence in a subject domain” (NRC, 2001, p. 44)

“A cognitive theory of how people develop competence in a domain” (NRC, 2001, p. 51)

“A cognitive model of learning” (NRC, 2001, p. 45)

“Ideally, then, a model of learning will also provide a developmental perspective, laying out one or more typical progressions from novice levels toward competence and then expertise, identifying milestones or landmark performances along the way” (NRC, 2001, p. 182).
Cognition vertex: Recovery

<table>
<thead>
<tr>
<th>Instrument Descriptors</th>
<th>PISA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>“reading literacy”</td>
</tr>
<tr>
<td>Sub-domain</td>
<td>processes</td>
</tr>
<tr>
<td>Sub-domain</td>
<td>text formats and types (continuous &amp; non-continuous)</td>
</tr>
<tr>
<td>Sub-domain</td>
<td>purpose of text</td>
</tr>
<tr>
<td>Taxonomy</td>
<td>NA</td>
</tr>
</tbody>
</table>
PISA’s Original Theory of Cognition: “Reading Literacy” as Processes

- **Retrieving information**
  - initial reading, forming a first impression before choosing to go into more detail. **Identifying** the main topic or message or through identifying the general purpose or use of the text, to describe the main character, setting or milieu of a story (OECD, 1999, p. 30)

- **Reflecting on the form of a text**
  - requires a high level of **metacognitive ability**. Readers must monitor their own thinking and reaction to a text. **Providing evidence** or arguments from outside the text, assessing the relevance of particular pieces of information or evidence, or **drawing comparisons** with moral or aesthetic rules (standards) (OECD, 1999, p. 32)

- **Reflecting on the content of a text**
  - comparing and contrasting information, drawing inferences, and identifying and listing supporting evidence (OECD, 1999, p. 31)

- **Forming a broad general understanding**
  - readers must scan and search the text, and **locate and select** relevant information (OECD, 1999, p. 31)

- **Developing an interpretation**
  - to stand apart from the text, consider it objectively, and **evaluate** its quality and appropriateness. These tasks include critical evaluation, and appreciation of the impact of such **textual features** as irony, humour and logical organisation. **Knowledge of text structure** play an important role… (OECD, 1999, p. 32)
PISA’s Original Theory of Cognition: “Reading Literacy” as formats and types of text

- Continuous
  - Description
  - Narration
  - Exposition
  - Argumentation
  - Instruction

- Non-continuous
  - Forms
  - Calls and advertisements
  - Charts and graphs
  - Tables and matrices
  - Maps
  - Diagrams

OECD, 1999
PISA’s Latest Theory of Cognition: “Reading Literacy” as Processes

Fluency: “is the ease and efficiency of reading texts for understanding” (OECD, 2019, p. 33).

Tasks require “students to read a sentence and make a judgment of the plausibility of the sentence based on general knowledge or the internal logical consistency of the sentence. The measure takes into account both the accuracy of the student’s understanding of the text and the time it takes to read and respond” (OECD, 2019, p. 24).

“One’s ability … to set up task-relevant reading goals, to monitor progress toward these goals, and to self-regulate their goals and strategies throughout the activity” (OECD, 2019, p. 36)
Observation vertex

<table>
<thead>
<tr>
<th>Item format (%)</th>
<th>Instrument</th>
<th>PISA (2000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed-choice</td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>Constructed Response</td>
<td></td>
<td>55</td>
</tr>
</tbody>
</table>
Table 2: Distribution of Reading Items by Text Structure and by Item Type

<table>
<thead>
<tr>
<th>Text Structure</th>
<th>Number of Items</th>
<th>Multiple-Choice</th>
<th>Complex Multiple-Choice</th>
<th>Closed Constructed-Response</th>
<th>Open Constructed-Response</th>
<th>Short Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous</td>
<td>89 (2)</td>
<td>42</td>
<td>3</td>
<td>3</td>
<td>34 (2)</td>
<td>7</td>
</tr>
<tr>
<td>Non-continuous</td>
<td>52 (7)</td>
<td>14 (2)</td>
<td>4 (1)</td>
<td>12</td>
<td>9 (1)</td>
<td>13 (3)</td>
</tr>
<tr>
<td>Total</td>
<td>141 (9)</td>
<td>56 (2)</td>
<td>7 (1)</td>
<td>15</td>
<td>43 (3)</td>
<td>20</td>
</tr>
</tbody>
</table>

Note: The numbers in parentheses indicate the number of items deleted after the main study analysis.

Table 3: Distribution of Reading Items by Type of Task (Process) and by Item Type

<table>
<thead>
<tr>
<th>Type of Task (Process)</th>
<th>Number of Items</th>
<th>Multiple-Choice</th>
<th>Complex Multiple-Choice</th>
<th>Closed Constructed-Response</th>
<th>Open Constructed-Response</th>
<th>Short Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interpreting</td>
<td>70 (5)</td>
<td>43 (2)</td>
<td>3 (1)</td>
<td>5</td>
<td>14 (2)</td>
<td>5</td>
</tr>
<tr>
<td>Reflecting</td>
<td>29</td>
<td>3</td>
<td>2</td>
<td>-</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>Retrieving Information</td>
<td>42 (4)</td>
<td>10</td>
<td>2</td>
<td>10</td>
<td>6 (1)</td>
<td>14 (3)</td>
</tr>
<tr>
<td>Total</td>
<td>141 (9)</td>
<td>56 (2)</td>
<td>7 (1)</td>
<td>15</td>
<td>43 (3)</td>
<td>20 (3)</td>
</tr>
</tbody>
</table>

Note: The numbers in parentheses indicate the number of items deleted after the main study analysis.

Adams & Wu, 2002
ACOL VOLUNTARY FLU IMMUNISATION PROGRAM

As you are no doubt aware, the flu can strike rapidly and extensively during winter. It can leave its victims ill for weeks.

The best way to fight the virus is to have a fit and healthy body. Daily exercise and a diet including plenty of fruit and vegetables are highly recommended to assist the immune system to fight this invading virus.

ACOL has decided to offer staff the opportunity to be immunised against the flu as an additional way to prevent this insidious virus from spreading amongst us. ACOL has arranged for a nurse to administer the immunisations at ACOL, during a half-day session in work hours in the week of May 17. This program is free and available to all members of staff.

Participation is voluntary. Staff taking up the option will be asked to sign a consent form indicating that they do not have any allergies, and that they understand they may experience minor side effects.

Medical advice indicates that the immunisation does not produce influenza. However, it may cause some side effects such as fatigue, mild fever and tenderness of the arm.

WHO SHOULD BE IMMUNISED?

Anyone interested in being protected against the virus.

This immunisation is especially recommended for people over the age of 65. But regardless of age, ANYONE who has a chronic debilitating disease, especially cardiac, pulmonary, bronchial or diabetic conditions.

In an office environment ALL staff are at risk of catching the flu.
PISA's Observation & Interpretation Vertex: Sample items, short task descriptions, associated scores

**Flu**

**ACOL Voluntary Flu Immunisation Program**
As you are no doubt aware, the flu can strike rapidly and extensively during winter. It can leave its victims ill for weeks.

The best way to fight the virus is to have a fit and healthy body. Daily exercise and a diet including plenty of fruit and vegetables are highly recommended to assist the immune system to fight this invading virus.

ACOL has decided to offer staff the opportunity to be immunised against the flu as an additional way to prevent this insidious virus from spreading amongst us. ACOL has arranged for a nurse to administer the immunisations at ACOL, during a half-day session in work hours in the week of May 17. This program is free and available to all members of staff.

Participation is voluntary. Staff taking up the option will be asked to sign a consent form indicating that they do not have any allergies, and that they understand they may experience minor side effects.

Medical advice indicates that the immunisation does not produce influenza. However, it may cause some side effects such as fatigue, mild fever and tenderness of the arm.

**Who Should Be Immunised?**
Anyone interested in being protected against the virus.

This immunisation is especially recommended for people over the age of 65. But regardless of age, ANYONE who has a chronic debilitating disease, especially cardiac, pulmonary, bronchial or diabetic conditions. In an office environment ALL staff are at risk of catching the flu.

---

Fiona McSweeny, the personnel officer at a company called ACOL, prepared the information sheet above for ACOL staff. Refer to the information sheet to answer the questions which follow.

**Question 6: FLU (R077Q2)**

**Reading Task:** Reading information

**Text Format:** Continuous

**Situation:** Occupational

Which one of the following describes a feature of the ACOL flu immunisation program?

A. Daily exercise classes will be run during the winter.

B. Immunisations will be given during working hours.

C. A small bonus will be offered to participants.

D. A nurse will give the injections.

**Scoring – Question 6**

Score 1: Answer B – Immunisations will be given during working hours.

Score 0: Other answers.

**Question 7: FLU (R077Q3)**

**Reading Task:** Reflection and evaluation

**Text Format:** Continuous

**Situation:** Occupational

We can talk about the content of a piece of writing (what it says).
We can talk about the style (the way it is presented).
Fiona wanted the style of this information sheet to be friendly and encouraging.
Do you think she succeeded?

Explain your answer by referring in detail to the layout, style of writing, pictures or other graphics.

---

**443 (Question 6)** This task requires students to locate explicitly stated information in a notice about an immunisation program in the workplace where competing or distracting information is present. (R077Q02)

---

**542 (Question 7)** This task requires students to evaluate the appropriateness of content in relation to the intended tone of a notice about immunisation. Readers need to draw on their understanding of what constitutes appropriate content for a particular purpose and audience. (R077Q03, score 1)

---

**583 (Question 7)** This task requires students to evaluate the appropriateness of formal text features in relation to the intended tone of a notice about immunisation. Readers need to draw on their understanding of what constitutes appropriate style for a particular purpose and audience. (R077Q03, score 2)

---

OECD, 2002
## Interpretation vertex

<table>
<thead>
<tr>
<th>Instrument</th>
<th>PISA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item format</td>
<td></td>
</tr>
<tr>
<td>fixed-choice</td>
<td>dichotomous (0,1)</td>
</tr>
<tr>
<td>constructed response</td>
<td>dichotomous (0,1)</td>
</tr>
<tr>
<td></td>
<td>polytomous (0,1,2)</td>
</tr>
</tbody>
</table>
For closed-response format items and for some open-constructed response items, the student has either chosen the designated correct answer or not: 1 (full credit) or 0 (no credit).

For more complex scoring of constructed response items, some answers, even though incomplete, indicate a higher level of reading literacy than inaccurate or incorrect answers, and receive partial credit: 2 (full credit) 1 (partial credit) or 0 (no credit).
The reading literacy scale is divided into first - 5, then (since 2015) 7 levels.

For 2018 items were designed to measure reading skill and understanding located at or below the current Level 1b.

<table>
<thead>
<tr>
<th>Level</th>
<th>What Students Can Do</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Readers at Level 6 typically can make multiple inferences, comparisons and contrasts that are both detailed and precise. They demonstrate a full and detailed understanding of one or more texts and may integrate information from more than one text. Tasks may require the reader to deal with unfamiliar ideas in the presence of prominent competing information, and to generate abstract categories for interpretations. Students can hypothesise about or critically evaluate a complex text on an unfamiliar topic, taking into account multiple criteria or perspectives and applying sophisticated understandings from beyond the text. A salient condition for accessing and retrieving tasks at this level is the precision of analysis and fine attention to detail that is incoherently in the texts.</td>
</tr>
<tr>
<td>5</td>
<td>At Level 5, readers can locate and organise several pieces of deeply embedded information, infering which information in the text is relevant. Reflective tasks require critical evaluation or hypothesis-making, drawing on specialised knowledge. Both interpreting and reflecting tasks require a full and detailed understanding of a text whose content or form is unfamiliar. For all aspects of reading, tasks at this level typically involve dealing with concepts that are contrary to expectations.</td>
</tr>
<tr>
<td>4</td>
<td>At Level 4, readers can locate and organise several pieces of embedded information. They can also interpret the nuances of language in a section of text by taking into account the text as a whole. In other interpreting tasks, students demonstrate understanding and application of categories in an unfamiliar context. In addition, students at this level can use formal or public knowledge to hypothesise about or critically evaluate a text. Readers must demonstrate an accurate understanding of long or complex texts whose content or form may be unfamiliar.</td>
</tr>
<tr>
<td>3</td>
<td>Readers at Level 3 can locate, and in some cases recognise the relationship between, several pieces of information that must meet multiple conditions. They can also integrate several parts of a text in order to identify a main idea, understand a relationship or construe the meaning of a word or phrase. They need to take into account many features in comparing, contrasting or categorising. Often the required information is not prominent or there is much competing information; or there are other text obstacles, such as ideas that are contrary to expectations or negatively worded. Reflecting tasks at this level may require connections, comparisons, and explanations, or they may require the reader to evaluate a feature of the text. Some reflecting tasks require readers to demonstrate a fine understanding of the text in relation to familiar, everyday knowledge. Other tasks do not require detailed text comprehension but require the reader to draw on less common knowledge.</td>
</tr>
<tr>
<td>2</td>
<td>Readers at Level 2 can locate one or more pieces of information, which may need to be inferred and may need to meet several conditions. They can recognize the main idea in a text, understand relationships, or construe meaning within a limited part of the text when the information is not prominent and the reader must make low-level inferences. Tasks at this level may involve comparisons or contrasts based on a single feature in the text. Typical reflecting tasks at this level require readers to make a comparison or several connections between the text and outside knowledge, by drawing on personal experience and attitudes.</td>
</tr>
<tr>
<td>1a</td>
<td>Readers at Level 1a can locate one or more independent pieces of explicitly stated information; they can recognise the main theme or author’s purpose in a text about a familiar topic, or make a simple connection between information in the text and common, everyday knowledge. Typically, the required information in the text is prominent and there is little, if any, competing information. The student is explicitly directed to consider relevant factors in the task and in the text.</td>
</tr>
<tr>
<td>1b</td>
<td>Readers at Level 1b can locate a single piece of explicitly stated information in a prominent position in a short, syntactically simple text with a familiar context and text type, such as a narrative or a simple list. Texts in Level 1b tasks typically provide support to the reader, such as repetition of information, pictures or familiar symbols. There is minimal competing information. Level 1b readers can interpret texts by making simple connections between adjacent pieces of information.</td>
</tr>
</tbody>
</table>
Next steps in our research: Validity and Reliability Checks on Score Interpretation

Evidence for
- content
  - panels
  - blueprints (Standards, DOK)
  - task demand analysis
- responses processes
  - think alouds/cognitive labs
  - exit surveys
- internal structure
  - dimensionality
- relations to external variables
- consequences

Evidence for
- internal consistency
- alternate forms
- test retest
- rater
  - intra
  - inter
Consequential Validity: The untold story of the Reading Wars

Our Professional Testing Standards warn us to examine: “the outcomes, intended and unintended, of using tests in particular ways in certain contexts and with certain populations” (p. 217).

We must examine potential implications for reading assessment “technologies” and “instruments” to inflict collateral damage on various populations.

Who is responsible for “testing events” that shape curriculum, teaching practices, and learning experiences? (The victors, the vanquished, or the “innocent ‘scientific-minded’ by-standers”).