

A Success Story from Kentucky

Richard Hill

CCSSO 30th Annual Conference on
Large-Scale Assessment

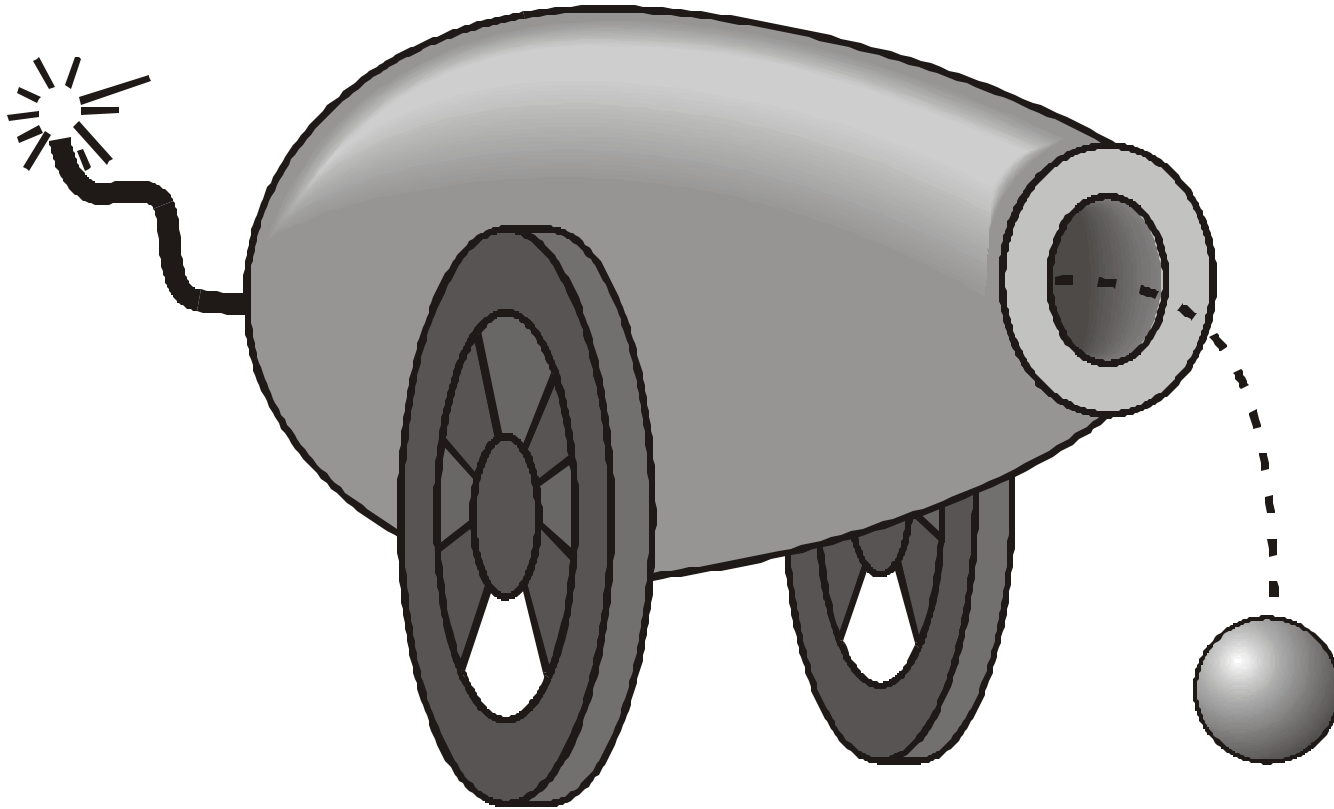
June 27, 2000

What We Thought We Were Doing



Hey, Claude, let's build us a performance assessment!

What We Actually Did



Sure, Maude. How about adding two short-answer questions to our 60-item multiple-choice test?

The Original Vision

- Teachers know what it is students are supposed to be able to do to reach achievement targets
- Active learners constantly produce artifacts that show what they know and can do (and what they don't yet know or can do)
- Teachers use these artifacts to plan new learning
- Teachers know at all times the achievement levels of all their students

The Reality—Part 1

- Teachers didn't change much
- Therefore, they didn't know their students' achievement levels
- Therefore, the external assessment had to collect more and more information
- Therefore, less efficient means of assessment fell by the wayside

The Reality—Part 2

- Teachers didn't change much
- Therefore, system needed to detect small amounts of change in performance
- Therefore, many misclassifications

The Success Story—Part 1

- Rescoring of 1992-93 writing portfolios
- Two samples:
 - ◆ Random
 - ◆ Purposeful (audit)

1992-93 Rescoring Results— Grade 4

<i>Score</i>	<i>All Schools</i>
<i>Original</i>	<i>32.7</i>
<i>Corrected</i>	<i>13.3</i>
<i>Difference</i>	<i>19.4</i>

1992-93 Rescoring Results— Grade 4

<i>Score</i>	<i>All Schools</i>	<i>Audit Schools</i>
<i>Original</i>	32.7	64.0
<i>Corrected</i>	13.3	19.6
<i>Difference</i>	19.4	44.4

Reaction

- Intent to correct scores
- Results sent to audit schools
- Protest
 - ◆ We rescored them wrong
 - ◆ The audit schools were being picked on

Regional Meetings

- Rapid change in understanding
 - ◆ Two categories of scoring elements
 - ◆ Long and neat does not equal Proficient
- Changes to training
 - ◆ Controlled by two people
 - ◆ Revised training materials
 - ◆ “high-end” portfolios
 - ◆ Practice materials and self-test

Two Years' Results Compared

	1992-93		1993-94
<i>Score</i>	<i>All Schools</i>	<i>Audit Schools</i>	<i>All Schools</i>
<i>Original</i>	32.7	64.0	37.5
<i>Corrected</i>	13.3	19.6	28.5
<i>Difference</i>	19.4	44.4	9.0

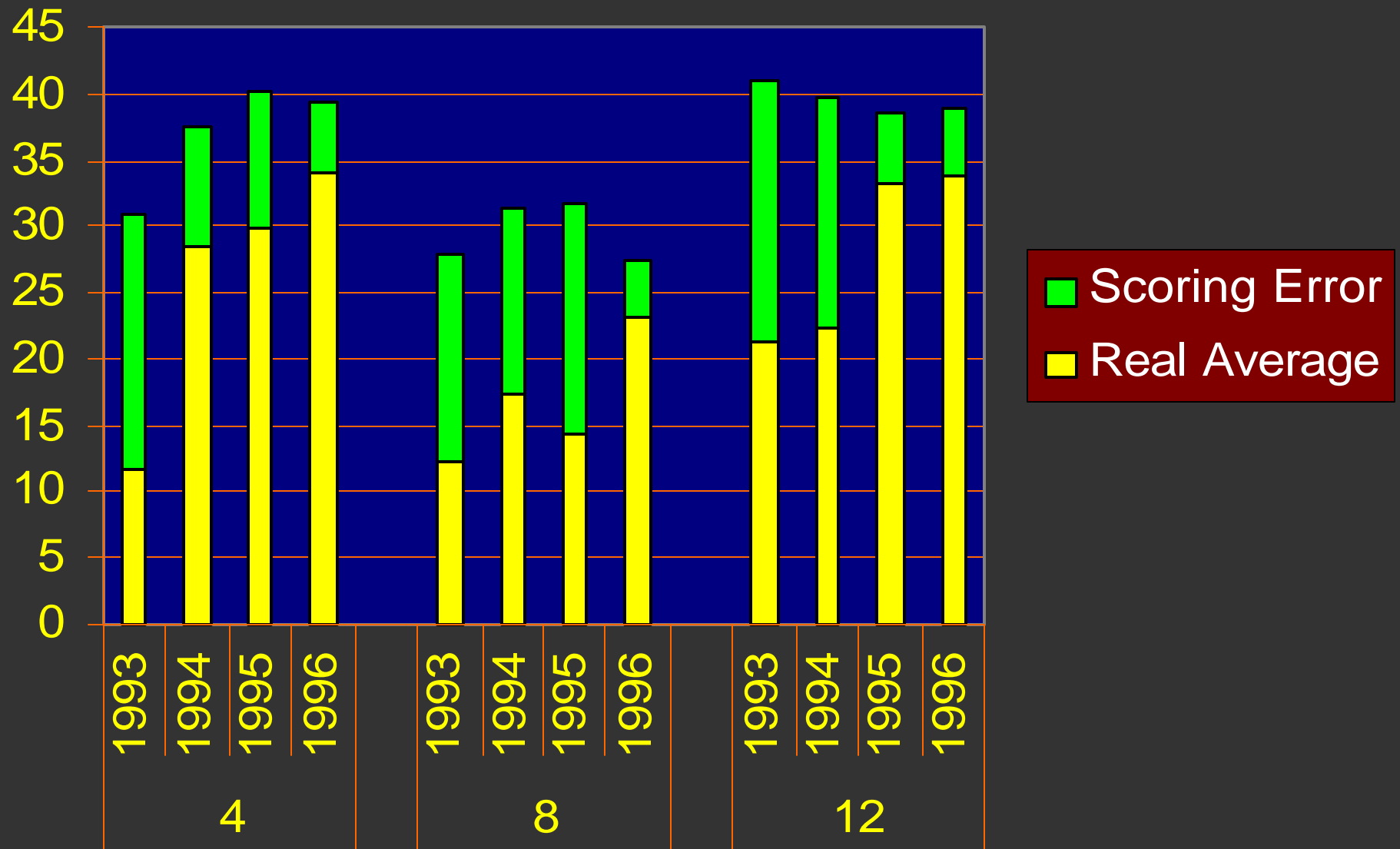
Two Years' Results Compared

	<i>1992-93</i>		<i>1993-94</i>	
<i>Score</i>	<i>All Schools</i>	<i>Audit Schools</i>	<i>All Schools</i>	<i>Audit Schools</i>
<i>Original</i>	<i>32.7</i>	<i>64.0</i>	<i>37.5</i>	<i>40.4</i>
<i>Corrected</i>	<i>13.3</i>	<i>19.6</i>	<i>28.5</i>	<i>37.4</i>
<i>Difference</i>	<i>19.4</i>	<i>44.4</i>	<i>9.0</i>	<i>3.0</i>

The Success Story—Part 2

- Realized feedback was critical
- Rescoring of a sample from every school
 - ◆ Half in 1994, remainder in 1995
 - ◆ Feedback on scoring accuracy
 - ◆ Feedback on next steps for improved writing
- Final rescoring study in 1996

Results, 1993-1996



Why Writing—and Writing Only?

- Writing is special
 - ◆ Natural production
 - ◆ More conceptual
 - ◆ Agreement about what quality writing is
- Extensive training already completed
- Accountability
- Resources