# Real-Time Data Use: Online Diagnostic Assessment System Developed to Help Teachers Improve Current Instruction in Grades 3 to 8

## National Center for Education Statistics
## 24th Annual Management Information Systems (MIS) Conference
## February 23 - 25, 2011
## Austin, Texas

**Real-time diagnostic instruments to identify
"why" each student is "low"
in Math and Reading Comprehension**

aka: **Error Identification assessments  (EIa)**

at 2009 American Educational Research Association

Alvin F. Larson, Ph.D.

al.larson@meriden.k12.ct.us

# A **Different** Type of "Assessment"
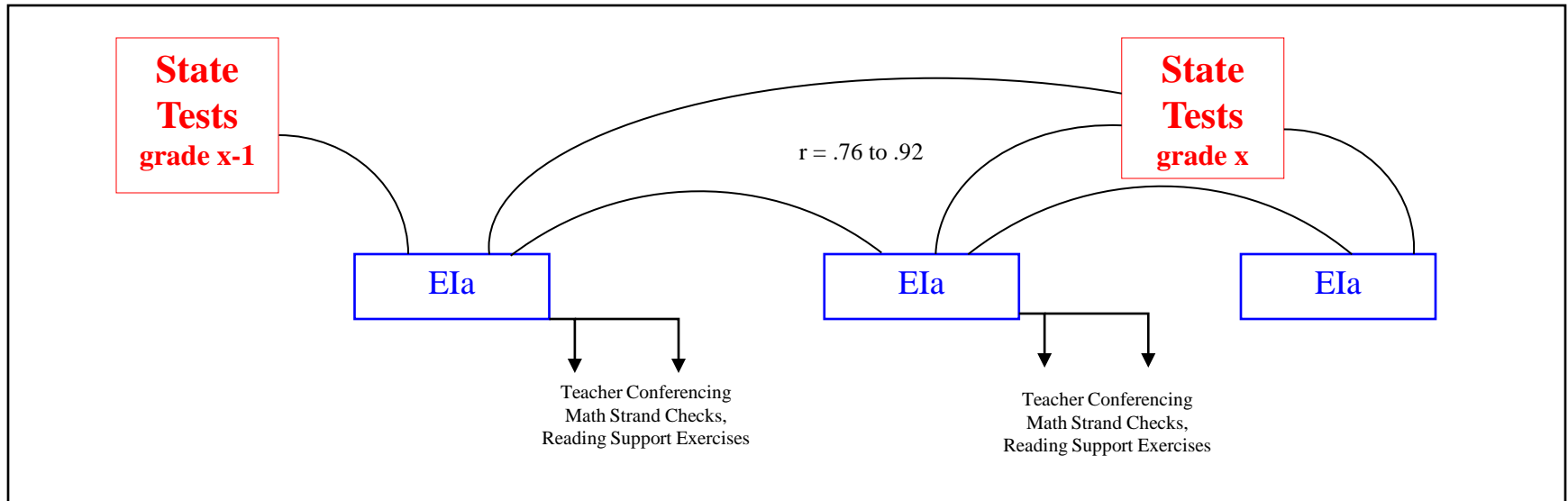
**State and "Standardized" Tests of the last century**

- designed to improve instruction by refining a rigid curriculum for the next year by providing reports **for** "decisionmakers"

- reports:

    - ability estimates (scale scores)

    - performance levels (evaluative)

    - mastery and raw score by "strand"

    - results reported 4 to 5 months after testing
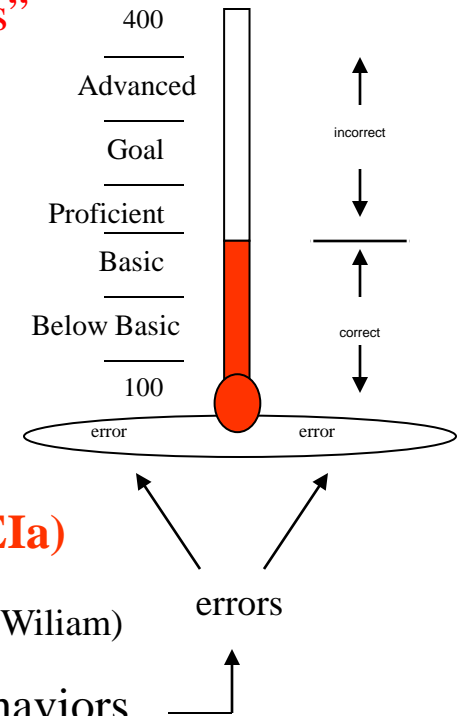
**Online Diagnostic: Error Identification assessments (EIa)**

- immediate non-evaluative feedback  (Butler, Black & Wiliam)

- reports student misconceptions and inefficient behaviors

    -  requires a "new" error vocabulary for Reading

- **designed** *for* **teachers** to improve their flexible *current* instruction that address student needs now



400

Advanced

Goal

Proficient

Basic

Below Basic

100

incorrect

correct

error          error

errors

# Purpose: The Integration of Instruction and Assessment [1]



- Calculated Validity and Reliability using professional statistical software
  - High Validity with CMT  (Math r = .82 to .92; Reading r = .76 to .86)
  - High Reliability (Math alpha = .88 to .94; Reading alpha = .80 to .89)
  - Point-biserials to help guide assessment development since Spring 2004

- Identify student misconceptions to inform current instruction
  - Theoretical support from Cognitive Diagnostic Assessments

- What teachers do: using the actual EIa items, teachers conference with students based on student misconceptions and adjust current instruction to meet individual student needs.

[1] *Test Theory for a New Generation of Tests* (Snow & Lohman, 1993)

# Demographics of
# Meriden Public Schools
# Meriden, CT

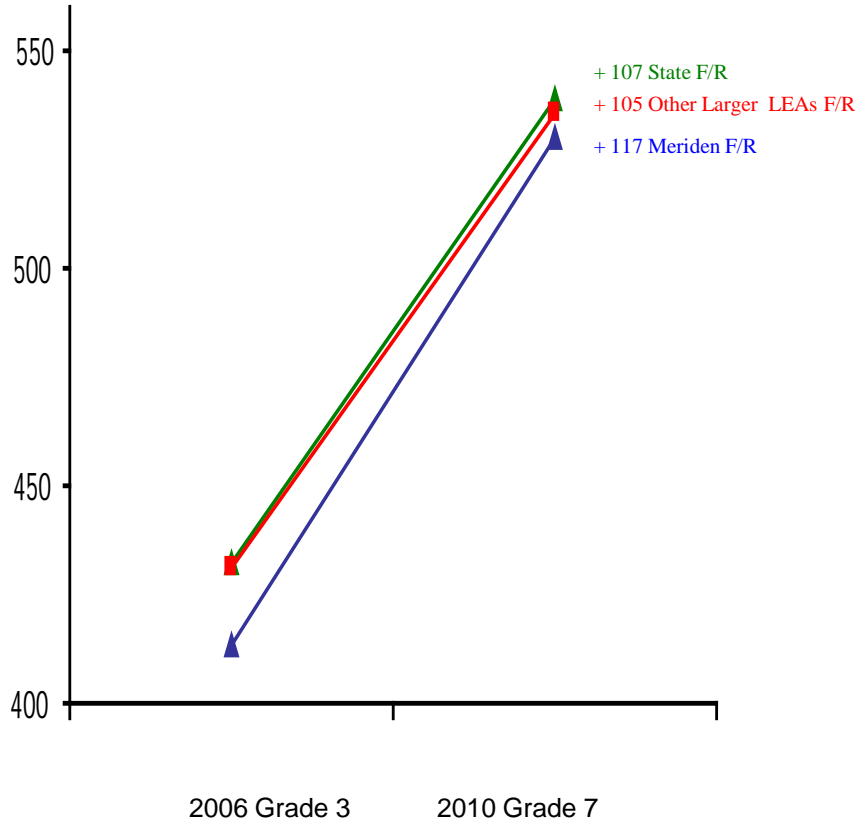| | |
|---|---|
| **40.2 %** | **White** |
| **43.8 %** | **Hispanic** |
| **13.5 %** | **Black** |
| **66.7 %** | **Eligible for Free/Reduced Meals** |
| | |
| Each year: Increasing minority and increasing poverty | |

Percent of Students At or Above Proficiency by Grade Level and Academic Area
State Tests 2006 to 2010
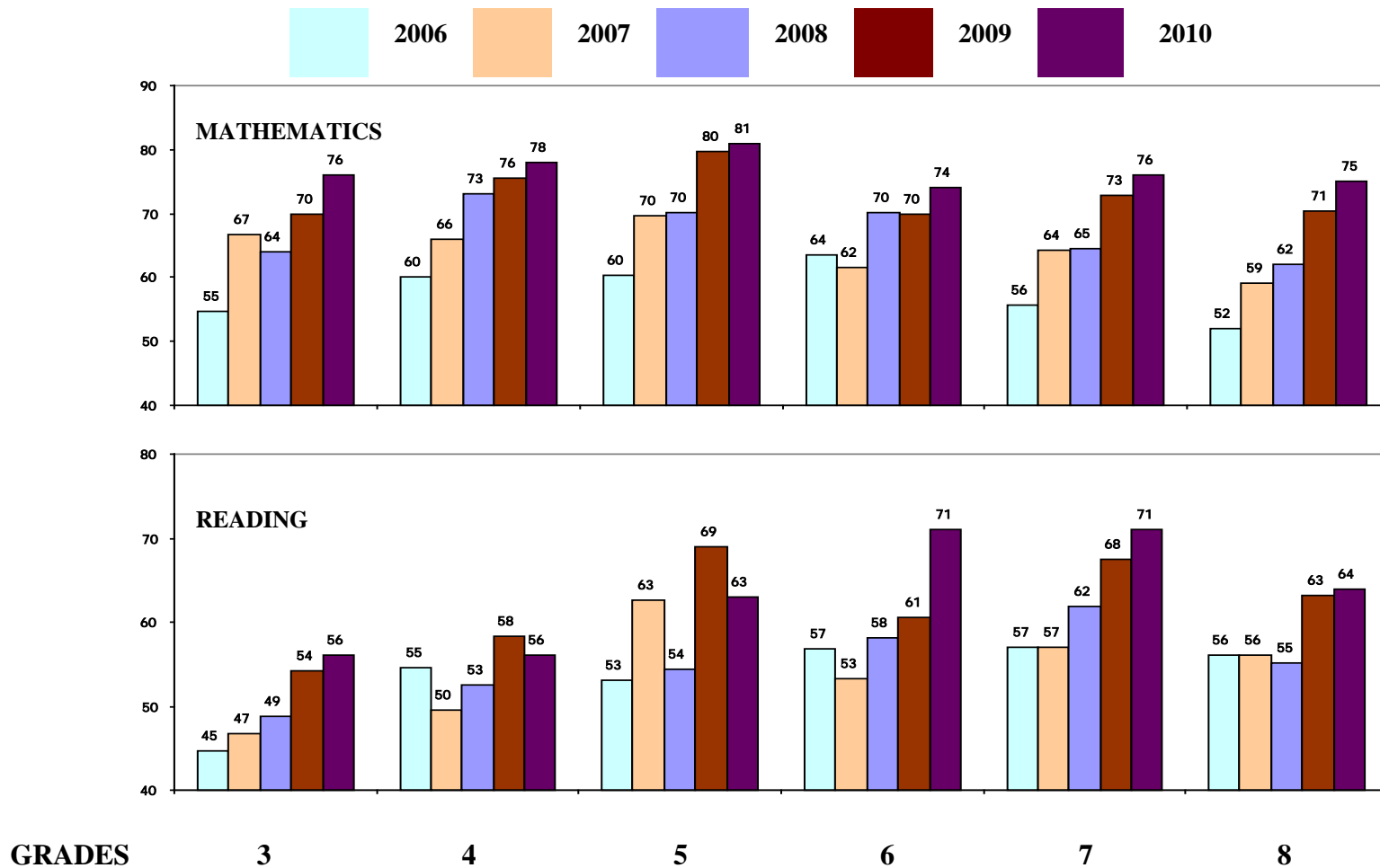
# "CLOSING THE SOCIOECONOMIC STATUS GAP"
## Five Year Matched Vertical Score Growth



MATHEMATICS

+ 107 State F/R
+ 105 Other Larger LEAs F/R
+ 117 Meriden F/R

550

500

450

400

2006 Grade 3          2010 Grade 7

READING

+96 State F/R
+ 94 Other Larger LEAs F/R
+ 102 Meriden F/R

510

460

410

360

2007 Grade 3          2010 Grade 7

6
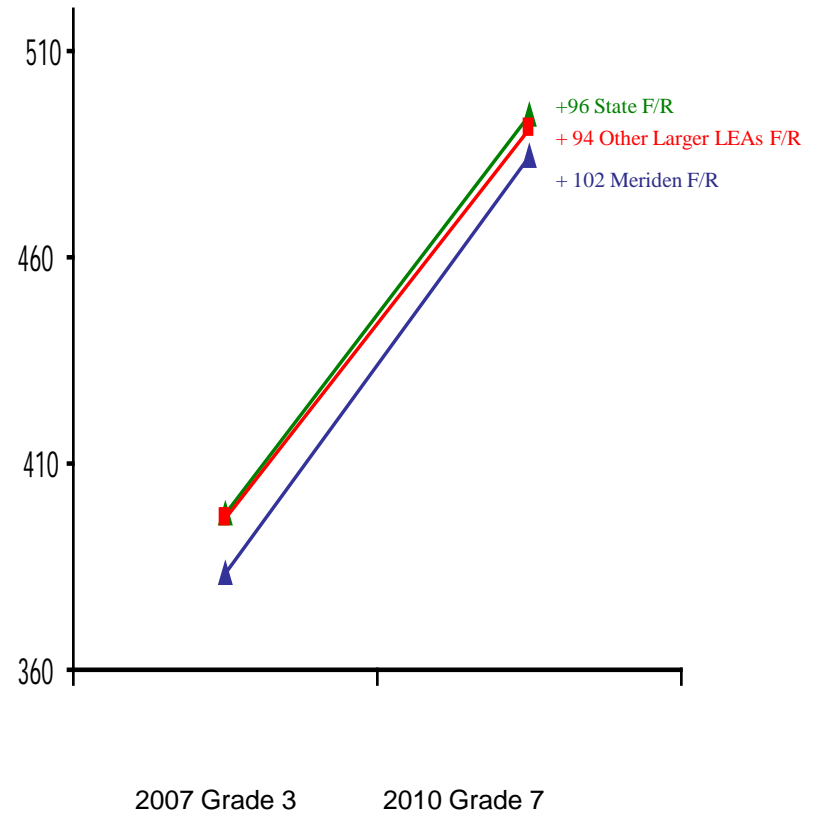
# LEA Purpose and Method

**Develop a district-wide assessment system that is meaningful to teachers in Math and Reading for grades 2-9 (10): (3 or 4 administrations per year)**

- **Theory from Cognitive Diagnostic Assessments, multiple-choice items were constructed by *designing* foils/distractors that mimic typical student cognitive processing errors**

- **Utilize Math/Reading error vocabulary to report error descriptions teachers understand**

Teachers need a finer "grain size" (NAEP and ETS)

# To be more "diagnostic," state's report strands

- Math Strand 1:  Place Value (45% Mastery)
  (about 4 types of questions)
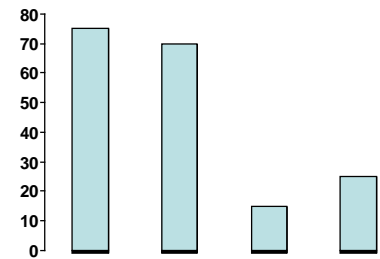
(4+ sub-strands)

Teachers need a finer "grain size" (NAEP and ETS)

# The Cognitive Task and Error Description in Math

MC Cognitive Task: evaluation of the differences between foils

$$\frac{1}{2} + \frac{1}{4} = \quad ?$$

A  1/6
B  2/4
C  2/6
D  3/4
E  1/8

# The Cognitive Task and Error Description in Math

MC Cognitive Task: evaluation of the differences between foils

$$\frac{1}{2} + \frac{1}{4} = \frac{2}{6}$$

A  1/6
B  2/4
C  2/6
D  3/4
E  1/8

Error Description: "adding both numerator and denominator"

a "finer grain size" that is immediate and meaningful
to both teachers and students

# A Finer Grain Size:

## Sample of Other Cognitive Error Descriptions in Math

### Four times a year x 4 foils x 45 items x 7 grades = 5,040 error descriptions

#8 = ' '     answer8='absent or left blank' .   /* strand 3 'Equivalent Fract, Decimal & Percents'.
#8 = 'A'     answer8='error: word prob; chose 1:6 vs. 1:3'.
#8 = 'B'     answer8='error: word prob; reversed ratio, chose 3:1 vs 1:3'.
#8 = 'C'      answer8=**'correct: word prob; 3:9 is the same as 1/3'.**
#8 = 'D'     answer8='error: word prob; chose 3:12 (1:4) ratio vs. 1:3'.
#8 = 'E'     answer8='error: word prob; chose 2:3 ratio vs. 1:3'.

#9 = ' '     answer9='absent or left blank' .   /* strand 4 'Order, Magnitude, and Rounding of Numbers'.
#9 = 'A'     answer9='error: ordering from table: selected 3rd place'.
#9 = 'B'     answer9='error: ordering from table: selected 1st place'.
#9 = 'C'     answer9='error: ordering from table: selected 2nd place'.
#9 = 'D'     answer9=**'correct: ordering from table: found 4 th place ordering from G to L'.**
#9 = 'E'     answer9='error: ordering from table: selected 5th place'.

#10 = ' '     answer10='absent or left blank' .   /* strand 5 'Models for Operations (one item)'.
#10 = 'A'     answer10=**'correct: word prob; chose correct number sentence ((6+4)X$5.00) for situation'.**
#10 = 'B'     answer10='error: word prob; divided instead of multiplying'.
#10 = 'C'     answer10='error: word prob; divided instead of multiplying'.
#10 = 'D'     answer10='error: word prob; added all data vs. adding 6 & 4, then multiplying'.
#10 = 'E'     answer10='error: word prob; subtracted instead of multiplying'.

# Using Diagnostic Errors to Improve Math Instruction

- Look for common errors across items…
  - A basic edict of cognitive diagnostic assessments

- Consider prerequisite skills and concepts needed…
  - Determine how to teach to the identified error

- Individual, small group, or whole group remediation of specific errors
  - Some flexibility of the delivered curriculum (for many items, mostly one typical error)

# Reading is Different from Math

## Both Require Thinking but…

- **You can "see" math errors**

    **…but not reading inferential thinking errors**


- **Math has a limited symbol domain (0, 1, 2, 3, 4, 5, 6, 7, 8, 9); Reading Comprehension symbol domain includes words, sentences, and inferred meaning on widely varied topics; Math has proofs, not inferences.**


- **Math teachers are trained in an error vocabulary that is aligned with misconceptions**

    **…reading literature emphasize strategies, not errors**


- **Math materials (text books) are very similar in both format and rigor to state and federal high stakes tests… but not reading**

# Disconnect: Testing vs. Teacher Training/Experience

**High Stakes Reading Tests are Inferential (higher level thinking)**
- Mostly Multiple Choice,  Items are difficult (p-values .3 to .7)

**Teacher Experience with Purchased Materials**
- too few multiple-choice and too easy (p-values .7 to .9)
- too literal (lower level thinking)
- teaching methods emphasize constructed-response items that are often too accepting, and without a rubric

**Teacher Training and Vocabulary**
- Metacognition (awareness of your reading comprehension) and fix-up strategies: look back to clarify, predict, author's purpose, main idea, activate background knowledge, etc. (Based upon Literature and NAEP Contexts)

**Need to develop a meaningful error vocabulary in Reading**
- Meaningful to teachers, based on their reading vocabulary

# For high stakes testing: What must students be able to do?

**Cognitive Model of Task Performance for
Reading Comprehension Multiple-choice Items**

**High Scoring versus Low Scoring Students**

**High scoring students will spend time and effort to think and**
- **be metacognitively aware;**
- **"look back" to clarify/re-read (good readers); and**
- **critical thinking: evaluate differences between foils**

**Low scoring students will**
- **be metacognitively less aware** . . . . . . . . . . . . (identified error: "X")
- **Tend not to look back; and** . . . . . . . . . . . . . (identified error: "T")
- **Tend not to evaluate…well** . . . . . . . . . . . . . . (identified error: "R")

# Reading Errors that are Meaningful to Teachers

**Each EIa foil is coded (pre-online coding 2004 to 2010):**

| CODE | ERROR in THINKING |
|------|-------------------|
| (X) | No-Support ("answer grabbing") |
| (T) | Text Matching ("look back" for a literal answer) |
| (R) | Related (good but not the "BEST" = "tricky") |
| (O) | Opposite |
| (A) | Anaphoric Pronoun Referent |
| (s) | Skipped Items |

# Text Matching (T) Did Not "Look Back" and No Support (X) Answer Grabbing Foils

## 58 Original Passages

| Exhibit 1 (Teacher Copy) |
|---|
| **JOHN ROCK** |

| Error Codes | Foils | Comment |
|---|---|---|
| Item 65 | Why did the author write **paragraph 4**? | (is the student re-reading?) |
| **T** | A. To show how he became a dentist. | (he was a "dentist" in paragraph 3) |
| **T** | B. To show John was a teacher. | (he was a "teacher" in paragraph 2) |
| **C** | C. To show John was a hard worker who wanted to help black people. | (inference) |
| **T** | D. To show John was the first black lawyer to be recognized by the Supreme Court. | (this was described in paragraph 5) |
| **X** | E. To show John moved to the South during the Civil War. | (he did not move to the South) |
| ↑ | ↑ | ↑ |
| **Error Codes** | **Foils** | **Comment** |

The purpose is not to get the right answer, but to discuss the errors and have the student improve his/her thinking and foster good reading behaviors.

# The Carefully Crafted … Related Foil:
## requires students to <u>evaluate</u> subtle differences between foils

| Exhibit 2 (Teacher Copy) |
| :---: |
| **CLIMBING HIGH** |

67      Which statement **BEST** describes what the passage is **MOSTLY** about?

| | | |
|---|---|---|
| **R** | A. Annie was a woman who took many risks. | *(she did take risks)* |
| **C** | B. Annie was a person who wasn't afraid to follow her dreams. | *(implied, best answer)* |
| **R** | C. Women can be mountain climbers. | *(she did)* |
| **T** | D. Most mountains can be climbed. | *(text matching)* |
| **T** | E. Most people never climb mountains. | *(text matching)* |

## The related foil, as a constructed-response, would be an
## acceptable <u>summative</u> answer

The purpose is not to get the right answer, but to discuss the errors and have the student improve his/her thinking and foster good reading behaviors.

# Sample Summary Reading Error Identification Report for a Teacher (Transitioning from Past to Future)



Error Identification in Reading: Sept 2010 - Grade 4

| | Reading Comprehension Summary | | | | | | | | click on the column header to sort | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Last | First | Level | ✓ | ♻ | 👁 | ☞ | ◐ | ∩ | A | B | D |
| | Nickolas | Proficient | 10 | 5 | 2 | 2 | 1 | 0 | 50% | 67% | 25% |
| | Savannah | Advanced | 17 | 3 | 0 | 0 | 0 | 0 | 90% | 83% | 75% |
| | Autumn | Below Basic | 3 | 7 | 6 | 3 | 1 | 0 | 10% | 33% | 0% |
| | Jordan | Advanced | 18 | 1 | 0 | 1 | 0 | 0 | 90% | 83% | 100% |
| | Alani | Proficient | 11 | 4 | 3 | 2 | 0 | 0 | 50% | 50% | 75% |
| | Anthony | Advanced | 18 | 1 | 0 | 1 | 0 | 0 | 90% | 83% | 100% |
| | Isaac | Below Basic | 5 | 4 | 7 | 4 | 0 | 0 | 30% | 0% | 50% |
| | Mychael | Goal | 13 | 2 | 3 | 2 | 0 | 0 | 60% | 67% | 75% |
| | Ashley | Advanced | 17 | 1 | 1 | 1 | 0 | 0 | 90% | 67% | 100% |
| | Estefania | Proficient | 10 | 0 | 4 | 6 | 0 | 0 | 40% | 33% | 100% |

Navigation:
- Home »
- Printable Probes »
- Overall Summary »
- Reading Summary »
- Passage 1 »
- Passage 2 »
- Passage 3 »
- Open Ended Responses »

Key:
- ✓ Correct Answer
- ♻ Related Answer
- 👁 Text Matching
- ☞ Answer Grabbing
- ◐ Opposite
- ∩ Skipped
- ⓘ View Question

Strands A, B, D are based on NAEP Contexts, also see:
Zwick, R. (1987). Assessing the Dimensionality of NAEP Reading Data.
Journal of Educational Measurement, Winter 1987, Vol 24, No. 4. pp 293-308

19

# Sample Individual Reading Error Identification Report for a Teacher



(Students need to explain their reasoning)

The purpose is not to get the right answer, but to discuss the errors and have the student improve his/her thinking and foster good reading behaviors.

20

# Sample Summary  Math Error Identification Report for a Teacher (Transitioning from Past to Future)



Error Identification in Math: Sept 2010 - Grade 5

**Math Summary** — click on the column header to sort

| Last | First | Level | ✳ | ✓ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|------|-------|-------|---|---|---|---|---|---|---|---|---|---|---|----|----|
|  | Brianna | Below Basic |  | 25 |  |  |  |  |  |  |  | D | A | A | B |
|  | Sierra | Goal |  | 40 |  |  |  |  |  |  |  |  |  | C | B |
|  | Luis | Goal |  | 44 |  |  |  |  |  |  |  |  |  |  |  |
|  | Victor | Proficient |  | 36 |  | B |  |  |  |  | E | A | A | A | B |
|  | Kevin | Basic |  | 33 |  | B |  | A |  |  |  | E |  | A | B |
|  | Anthony | Goal |  | 45 |  |  |  |  | E |  |  |  |  |  |  |
|  | Jazmin | Basic |  | 32 |  |  |  | A |  |  |  |  |  | C | B |
|  | Brian | Below Basic |  | 22 |  | B |  | C | A |  |  | A | B | C | C |
|  | Courtney | Basic |  | 30 |  |  |  |  |  |  |  | D | C | C | C |
|  | Hector | - |  | - | - | - | - | - | - | - | - | - | - | - | - |

**Key:**
- ✳ Scale Score
- ✓ Correct Answer
- ∩ Skipped
- ⊙ View Question

*Mouse-over an icon to the right to see its meaning.*

Home »
Printable Probes »
Overall Summary »

(Students often make same cognitive error)

In Math, the purpose is to identify probable student errors and guide specific instruction/review.

21

# Sample Individual Math Error Identification Report for a Teacher (non-evaluative)



Home »
Printable Probes »

Overall Summary »

**Key:**
- Scale Score
- Correct Answer
- Skipped
- View Question

*Mouse-over an icon to the right to see its meaning.*

**Error Identification in Math: Sept 2010 - Grade 5**

**Kevin**

| Question # | Error Description |
|---|---|
| 1 | |
| 2 | B: error: word prob; added 100 instead of subtracting 100: (443 - 100 = 543) |
| 3 | |
| 4 | |
| 5 | A: error: pictorial rep of mixed #; used unshaded as numerator: (1 4/5 = 1 1/5) |
| 6 | |
| 7 | |
| 8 | E: error: number line; selected too much on scale: (place 0.4 ml on # line; = 0.9 ml) |
| 9 | |
| 10 | A: error: word prob; added vs subtracted (R has 48 & S has 71; how many more = 71 + 48) |
| 11 | B: error: word prob; subtracted vs added & incorrect order (78 and 53 more = 53 - 78) |
| 12 | C: error: word prob; multiplied rather than divided (42 in bunches of 7 = 42 x 7) |
| 13 | |

# The Integration of Instructional and Assessment Functions [1]

**Teachers must Adjust their Instruction (Curriculum Pacing of the Last Century) based on Identified Student Needs**

## The actual test items and foils are used as instructional aids

### Teachers conference with students (students explain their reasoning)

- an "internal view" with "retroactive verbal reports" (Leighton & Gierl; Norris; Gorin; 2007); "think alouds" (Davey, 1983);

- Help students get involved in their own learning by making their thinking visible to themselves, peers and teachers; and

- For the teaching of critical thinking (inference) and understanding of ideas in the text [foils] (Wells, 2000; Block, Gambrell & Pressley, 2002).

### Additional instructional/assessment materials in Reading

- Reading Comprehension Support Exercises

[1] *Test Theory for a New Generation of Tests* (Snow & Lohman, 1993)

# Reading Comprehension Support Exercise
## Salute to the Irish Brigade

Last week, despite the inclement weather, patriots turned out in numbers recently for the wreath-laying ceremonies at the Ninth Regiment Civil War monument in the Hill section of the city.  They paid tribute to the brave soldiers who lost their lives fighting for the "Ninth," made up mainly of Connecticut Irishmen, during the Civil War.

The Second Company Governor's Foot Guard Fife and Drum Corps provided music.  A piper from the Gaelic Highland Pipe Band played the bagpipes while a New Haven police officer and bugler, stood ready.
A reception was held following the ceremony at the New Haven Gaelic Club in East Haven.

In the author's "Salute to the Irish Brigade," who were the "patriots" who went outside in bad weather for the wreath-laying ceremonies?

| | | |
|---|---|---|
| *Text Matching* | A. | The Ninth Regiment of the Civil War |
| *Text Matching* | B. | Connecticut Irishmen during the Civil War |
| *Related* | C. | The Second Company Governor's Foot Guard |
| *Answer Grabbing* | D. | George Washington and the Continental Army |
| *Correct* | E. | The people watching and participating in the activities at the Hill section of the city last week. |

# Teacher Opinion of EIa:
### (only one implementation study with a 17% return rate)

➢ I use them to see what they are thinking … which errors seem to be used consistently - what they are "tricked on."  I turn it into a game-type activity: students vs. teacher (reading teacher, grades 2-5).

➢ If students are aware of the errors they make, determined students will change their behavior both in reading and testing (classroom teacher, grade 5).

➢ They now don't grab the first answer that they connect to, they take more time to evaluate and critique each choice (reading teacher, unknown grade).

➢ No support errors identifies students who clearly can't read on grade level or aren't taking it seriously (classroom teacher, grade 4).

# Teacher Opinion of EIa:

➢ I meet with students in small groups… Keeping the test booklets and handing them back to students is helpful in discussing why they chose their answers and it makes them accountable for their choices (classroom teacher, grade 4).

➢ Understanding their misconceptions is one thing – getting them to change the misconception is the difficult part (classroom teacher, grade 7).

➢ As we are working on it, [EIa post-conferencing] often students will "get" it when it is a clear error.  The related [versus the] correct answer does not come as easily during the explanation (classroom teacher, grade 4).
     (Related is an important foil for "ability estimates" or critical thinking, some students will need more reading experience and scaffolding)

The Error Identification component of Meriden District Assessments has been a successful, seven-year (paper/pencil) project* which addresses the feasibility of an LEA developing its own Cognitive Diagnostic Assessments for classroom use.**

These relatively short assessment tools (currently online 2010-2011):

- Explain "why" each student might be "low" *for* teachers
  (after the teacher–student conferencing with the actual test items),

- Guide adjustments to current instruction and review exercises
  (the integration of Instructional and Assessment functions), ***

- Are highly correlated with and have a similar level of difficulty
  as the required state assessments, and

- Are ready for dissemination to other LEAs through a
  free* online system capable of providing immediate reports.

---

* This "legacy" project was partially funded with Title I funds, so items are free, but there are technical support costs.

** Gierl & Leighton (2007). Directions of Future Research in Cognitive Diagnostic Assessment. In M. J. Gierl & J. P. Leighton (Eds.), *Cognitive diagnostic assessment for education theory and applications* (p. 349). New York, NY: Cambridge University Press.

*** Test Theory for a New Generation of Tests (Snow & Lohman, 1993).

# Web Solutions Pricing Models

- ## Application Service Provider Model

    » Hosting Price Per Student Per Year                                           $2.00
      *Includes 10 hrs per year of support services. (Extra blocks of time available)*

    » One Time Initial Setup of Student Data                               $750-$1,500
       *Depending upon provided format data*

- ## Setup Model

    » One Time Setup Agreement                                                  $15,000
      *Your server environment must include: MS Windows 2008 Server & SQL Server 2008*

    » One Time Initial Setup of Student Data                               $750-$1,500
      *Depending upon provided format data.*

    » Required Yearly Support Packages

| | hrs/month | rate/hour | yearly fee |
|---|---|---|---|
| Basic | 2 | $110 | $2,640 |
| Standard | 3 | $105 | $3,780 |
| Enhanced | 4 | $100 | $4,800 |