



New Knowledge about Writing

Richard Hill

The Center for Assessment

2002 Annual Large Scale
Assessment Conference



Greetings from Pennsylvania

- Thanks to PDE staff for supporting this study
- Thanks to DRC for successfully completing all the tasks



Background

- Chapter 4
- Student-level reporting for first time
- Moderate stakes for grade 11 students
- Concern for reliability of writing scores
- Concern about amount of time and money for scoring
 - Requirement that results be reported by standard



Three Goals

- Generalizability Study
 - More tasks needed?
 - Each additional task adds 10 weeks to reporting schedule and \$500,000 to cost
- Determine optimal adjudication rules
- Study to see whether scoring system produces subscores that provide unique information



Data Needed

- 1800 students respond to four prompts each
- Six studies of 300 students each
 - 1—Four narrative prompts
 - 2—Four informational prompts
 - 3—Four persuasive prompts
 - 4—Two narrative and two persuasive
 - 5—Two info. and two persuasive
 - 6—Two narrative and two info.



Scoring

- Each paper scored by 2 teams of 15 scorers
- 3 scorers from each team scores each paper

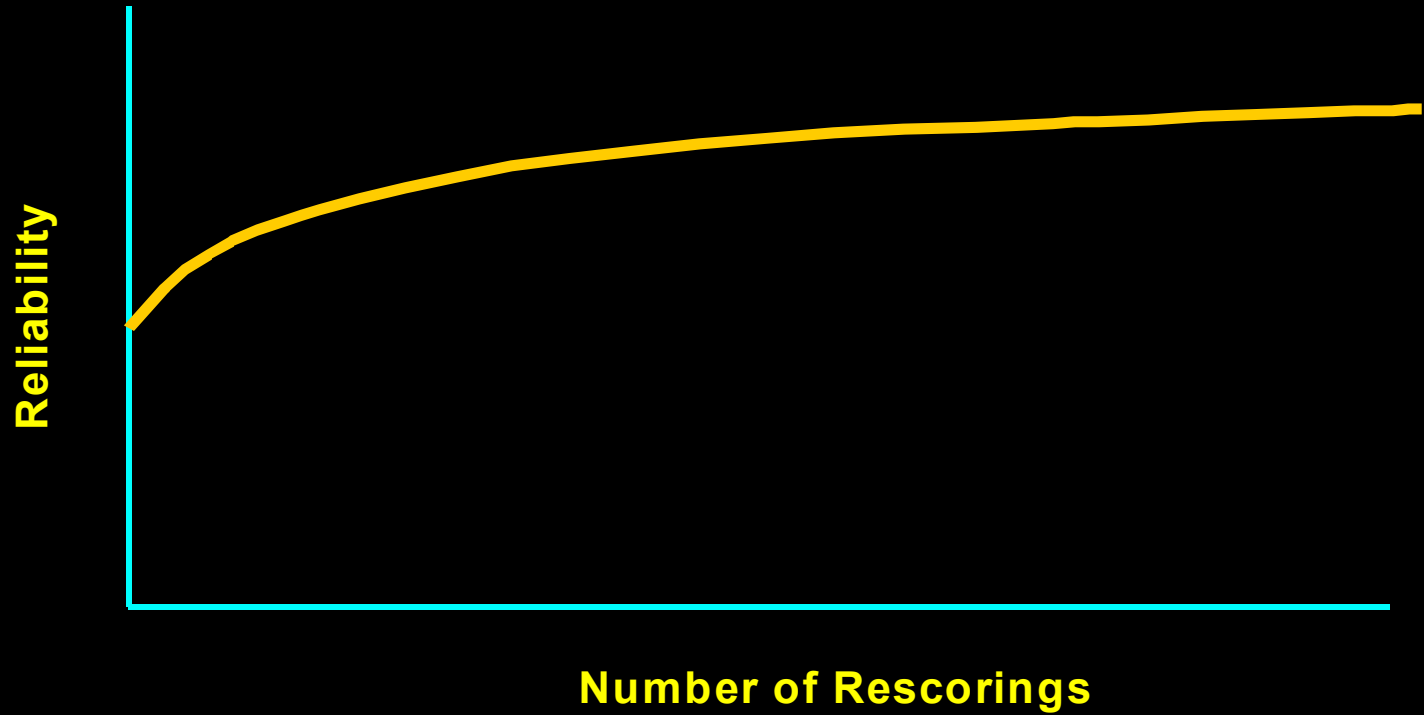


Scoring, cont'd.

- Each team consists of 3 scorers
- For a given adjudication rule
 - Score every paper by both teams using that rule
 - Determine the number of scorings
 - Determine the consistency between the two teams



Desired Data





Initial Look

	No Adjudication
% Additional Rescores	0
Correlation	.85
Ave. Diff.	2.6
% within 2	64



Initial Look

	No Adjudication	Adjudicate Any Discrepancy
% Additional Rescores	0	79
Correlation	.85	.85
Ave. Diff.	2.6	2.5
% within 2	64	68



Focus on Final Decision—One Paper

1. Adjudicate nothing
 - A. Both pass
 - B. Either passes
2. Adjudicate if passed by one but failed by another—pass if 3rd scorer passes
3. Adjudicate if within 4 points of passing—pass if total of two highest scores ≥ 25



Results for One Paper

	1A	1B	2	3
% Rescored	0	0	18	19
% Failed by Both Teams	33	18	26	18
% Passed by One Team but Failed by Other Team	18	14	13	12



Focus on Final Decision— Three Papers

- Adjudicate nothing
- Adjudicate any paper failed by one scorer but passed by second; take sum of two highest scores
- Adjudicate any paper within 4 of passing; take sum of two highest
- Adjudicate any paper of 11 or 12; take sum of two highest
- Adjudicate any total of 71-74; rescore all three papers and take sum of two highest scores on each



Results for Three Papers

	1	2	3	4	5
% Rescored	0	18	15	28	7
% Failed by Both Teams	21	19	20	19	18
% Passed by One Team but Failed by Other Team	9	9	8	9	8



Correlations among Ratings

	1	2	3	4	5
1	1.00	.80	.83	.72	.60
2		1.00	.80	.76	.57
3			1.00	.76	.63
4				1.00	.66
5					1.00



Same Scorer \ Different Scorer

	1	2	3	4	5
1	.62\.66	.64	.63	.59	.49
2	.66	.76\.70	.63	.60	.48
3	.58	.64	.62\.66	.59	.51
4	.50	.62	.55	.50\.64	.52
5	.49	.51	.53	.43	.64\.61



Same Scorer Twice

	1	2	3	4	5
1	.62	.66	.58	.50	.49
2	.96	.76	.64	.62	.51
3	.94	.93	.62	.55	.53
4	.90	1.00	.99	.50	.43
5	.78	.78	.84	.76	.64

Each Scorer Scores One Domain



	1	2	3	4	5
1	.51	.61	.58	.55	.45
2	.99	.74	.67	.63	.46
3	1.00	.96	.66	.59	.48
4	1.00	1.00	1.00	.50	.50
5	.81	.68	.76	.91	.61

Same Scorer, Same Student, Different Paper

	1	2	3	4	5
1	.49	.51	.49	.48	.43
2	.99	.54	.50	.50	.42
3	1.00	.97	.49	.48	.44
4	.96	.95	.96	.51	.45
5	.84	.79	.86	.87	.53

Different Scorer, Same Student, Different Paper

	1	2	3	4	5
1	.47	.49	.47	.47	.41
2	.99	.52	.48	.48	.41
3	1.00	.97	.47	.47	.42
4	.98	.95	.98	.49	.45
5	.85	.81	.88	.92	.49



Grade 6

	1	2	3	4	5
1	1.00	.76	.78	.73	.64
2		1.00	.81	.79	.66
3			1.00	.80	.70
4				1.00	.76
5					1.00



Grade 9

	1	2	3	4	5
1	1.00	.85	.84	.81	.80
2		1.00	.85	.82	.79
3			1.00	.84	.81
4				1.00	.87
5					1.00



Grade 6

	1	2	3	4	5
1	.38	.39	.41	.41	.42
2	.96	.43	.43	.44	.42
3	1.00	.99	.44	.46	.45
4	.98	.97	1.00	.48	.49
5	.93	.87	.92	.96	.54



Grade 9

	1	2	3	4	5
1	.46	.47	.47	.47	.47
2	1.00	.48	.48	.48	.48
3	.99	.99	.49	.49	.49
4	.98	.98	.99	.50	.50
5	.97	.97	.98	.99	.51



For info on data set, contact:

Lenny Lock, (717-787-4234)
llock@state.pa.us

Richard Hill
Center for Assessment

Rhill@nciea.org

www.nciea.org

