

# **Relationship between the Fountas & Pinnell Benchmark Assessment System and the Smarter Balanced ELA Assessments, Grades 3-5**



Washington Educational Research Association (WERA) Preconference  
December 2015 \* Seatac, WA

**Jack B. Monpas-Huber, Ph.D.**

Director of Assessment, Student Information & Highly Capable Program  
Marysville School District

**Brian Rick**

Director of Research and Assessment  
Bellingham School District

**Debra Hawkins, Ed.D.**

Director of Assessment and Research  
Issaquah School District

# About the Fountas and Pinnell Benchmark Assessment System

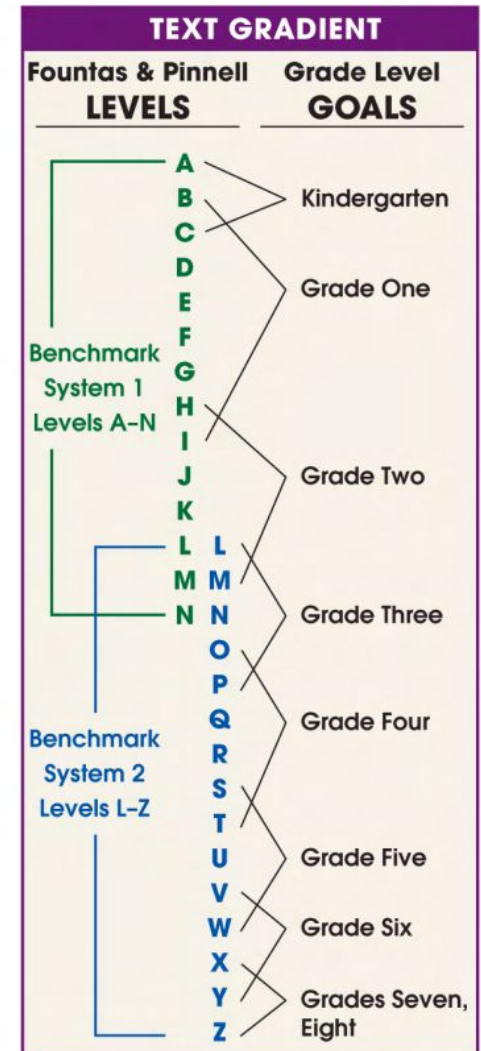
## “F & P Gradient of Text Difficulty”

Texts are examined according to 10 criteria and then assigned to levels A through Z

Goal of leveled text for students

- Read increasingly complex fiction and nonfiction
- Develop strategies to read independently

One-on-one assessment to determine independent and instructional reading levels and for placing students on the F & P Text Gradient™, A-Z



# About the Fountas and Pinnell Benchmark Assessment System

## The F & P assessment process

Student reads a leveled text

Teacher listens to reading, records rubric score from 0 to 3

Measures of accuracy, fluency, comprehension

Overall independent and instructional reading levels

Recording Form More Than a Pet • LEVEL 3 • NONFICTION

**Part Two: Comprehension Conversation**

Have a conversation with the student, noting the key understandings the student expresses. Use prompts as needed to stimulate discussion of understandings the student does not express. It is not necessary to use every prompt for each book. Score for evidence of all understandings expressed—with or without a prompt. Circle the number in the score column that reflects the level of understanding demonstrated.

*Teacher:* Talk about what you learned in this book.

**Comprehension Scoring Key**  
0 Reflects **unsatisfactory** understanding of the text. Either does not respond or talks off the topic.  
1 Reflects **limited** understanding of the text. Mentions a few facts or ideas but does not express the important information or ideas.  
2 Reflects **satisfactory** understanding of the text. Includes important information and ideas but neglects other key understandings.  
3 Reflects **excellent** understanding of the text. Includes almost all important information and main ideas.

Key Understandings	Prompts	Score
<b>Within the Text</b> There are two kinds of dogs that help people. Some dogs help people feel better and some dogs work. (May or may not use the terms therapy and service.) Dogs help people in many different ways. (Gives 2-3 examples such as: dogs cheer people up; they help their owners; they make sure the owner is safe; they get things for their owners.) <i>Note any additional understandings:</i>	What were the two kinds of dogs that you read about? What did you learn about therapy dogs? What did you learn about service dogs? What else did you learn about the two kinds of dogs?	0 1 2 3
<b>Beyond and About the Text</b> Dogs must be smart because they can learn to help people in different ways (or people train them). Dogs are good pets and they can also do a lot more things to help people. In this picture, the woman is blind and the dog is helping her get on a train. The heading helped me know what kind of dog I would be reading about. <i>Note any additional understandings:</i>	Why do you think dogs can be so helpful? Why do you think dogs are so important to people? Look at the photograph on page 13. What information does it give you? Look at page 4. How does the heading "Therapy Dogs" (point to it) help you?	0 1 2 3

**Guide to Total Score**  
4-3 Excellent Comprehension  
3 Satisfactory Comprehension  
2 Limited Comprehension  
0-1 Unsatisfactory Comprehension

**Subtotal Score:** \_\_\_\_/6  
**Add 1 for any additional understandings:** \_\_\_\_/1  
**Total Score:** \_\_\_\_/7

**Part Three: Writing About Reading (optional)**

Read the writing/drawing prompt below to the student. You can also cut the prompt on the dotted line and give it to the child. Specify the amount of time for the student to complete the task on a separate sheet of paper. (See Assessment Guide for more information.)

**Writing About Reading Scoring Key**  
0 Reflects **no** understanding of the text.  
1 Reflects **very limited** understanding of the text.  
2 Reflects **partial** understanding of the text.  
3 Reflects **excellent** understanding of the text.

Write about three ways dogs help people. You can draw a picture to go with your writing.

Fountas & Pinnell Benchmark Assessment System 1

# Methods

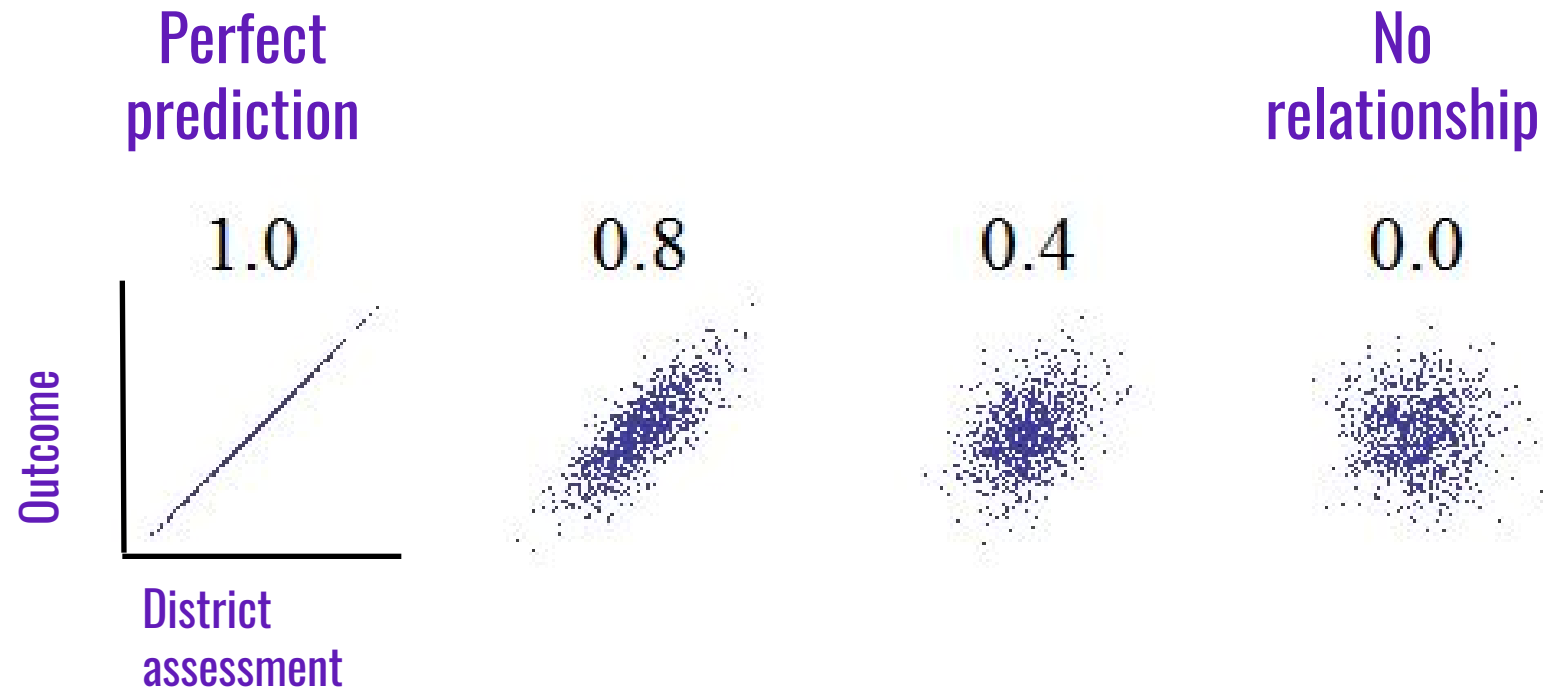
- Combined data from Bellingham, Issaquah and Marysville
- 2014-15
- Grades 3-5
- Only students who had three data points
  - Fall F & P level
  - SBA ELA scale score
  - Spring F & P level

# Methods

Sample sizes

Grade	Bellingham	Issaquah	Marysville
3	414	570	641
4	399	396	619
5	159	0	535

# Review: What does a correlation look like?



# Results: Pearson correlations to Smarter Balanced ELA

Grade	N	Fall F & P	Spring F & P
3	1,625	0.67	0.69
4	1,414	0.68	0.68
5	694	0.60	0.65

# Results: Diagnostic Accuracy of F & P

Grade	F & P Spring on grade level	Met Smarter Balanced standard		
		No	Yes	Total
3	No	395	36	431
	Yes	442	752	1,194
	Total	837	788	1,625
		52.8%	95.4%	
4	No	336	47	383
	Yes	357	674	1,031
	Total	693	721	1,414
		51.5%	93.5%	
5	No	273	65	338
	Yes	118	238	356
	Total	391	303	694
		30.2%	78.5%	

## Observations

Except for 5th grade...?

F & P is very sensitive to proficiency

Correctly classified of 90% of SBA-proficient students in 3rd and 4th

But also overpredicted proficiency by hundreds of students

Or did SBA under-classify proficient students?



# Implications for our work

## F & P is a valuable tool, enjoys a lot of support

- One-on-one; teachers hear students read
- Written records (i.e., low tech?)
- Diagnostic and actionable
- Not just an assessment, is part of an instructional system that provides clear instructional implications and resources

# Implications for our work

Always good when new data expands our understanding of assessment/s

- THREE reading assessments laying claim to grade level proficiency
- We now know that STAR is a stronger predictor of SBA than F & P . . . for what that's worth
- Correlations and criterion validity are new ways of evaluating the quality of our assessments

# Implications for our work

## These results sharpen thinking about our reading assessments

- More clearly distinguish screening from diagnosis as formative purposes
  - Let STAR accurately screen students into risk groups and measure growth
  - Let F & P provide diagnosis and guide intervention and instruction

# Implications for our work

But the aggregate F & P data raises questions for more validity studies:

- There is subjectivity in scoring F & P. This introduces rater variance. How to isolate that?
- F&P is ordinal measurement: levels are not equidistant. This limits what we can do (i.e., classical statistics based on interval data)
- F & P growth and STAR growth? Do they complement?
- How does F & P growth relate to SBA proficiency?

THANK  
YOU