

2008

Net Impact Analysis of Reading Recovery[®] in Kalamazoo Public Schools: Concept and Preliminary Findings

Randall W. Eberts

W.E. Upjohn Institute, eberts@upjohn.org

Wei-Jang Huang

W.E. Upjohn Institute

Citation

Eberts, Randall W., and Wei-Jang Huang. 2008. "Net Impact Analysis of Reading Recovery[®] in Kalamazoo Public Schools: Concept and Preliminary Findings." Prepared for Kalamazoo Community Foundation.

<https://research.upjohn.org/presentations/1>

This title is brought to you by the Upjohn Institute. For more information, please contact repository@upjohn.org.

Net Impact Analysis of Reading Recovery® in Kalamazoo Public Schools: *Concept and Preliminary Findings*

November 18, 2008

**Randall W. Eberts
Wei-Jang Huang**

W.E. Upjohn Institute for Employment Research

Purpose

- **Purpose:** To conduct a net impact analysis of the performance of KPS students who participated in Reading Recovery®.
- **Question:** Does Reading Recovery® make a difference in the reading ability of students who participate in the program? That is, do students in Reading Recovery® improve their reading ability more than if they had not participated in the program?
- **Net Impact Analysis:** Estimates the impact of a program by comparing the outcomes of participants with the outcomes of “identical” individuals who have not participated in the program.
 - For Reading Recovery®, if the outcome of participants is larger than the outcomes of those not participating in the program, and the difference is statistically significant, then it can be concluded with a high degree of confidence that Reading Recovery® has a positive net impact on the participants in the program.
- **Need for Net Impact Analysis:** If one wants to evaluate a program’s contribution in helping participants achieve their objectives, one must sort out the contribution of the program from the contribution of non-program factors.
 - For Reading Recovery®, this means sorting out the contribution of the program from the innate abilities of the students, teacher, school and home influences, and additional outside help
 - Without sorting out these extraneous effects, it is impossible to isolate the effect of the program from the effects of other factors

Reading Recovery®

- **Reading Recovery® is a literacy intervention for first grade students who are at risk of not becoming successful readers. It is a short-term (12 to 20 weeks), one-on-one tutorial designed to help first graders who are in the lowest 20th percentile of their class. Reading Recovery® teaches reading and writing strategies with the goal of getting students to achieve average classroom literacy proficiency. Students engage in reading and writing with a Reading Recovery® teacher who supports learning activities. Over time, the student gradually takes over the process of reading and becoming an independent reader.**
- **Reading Recovery® is comprised of a diagnostic survey, tutoring sessions, and teacher training.**
 - The diagnostic survey considers six measures: text reading level, letter identification, concepts about print, word test, writing vocabulary, and dictation.
- **Each Reading Recovery® tutorial is 30 minutes long and consists of a standard framework of activities designed by the trained teacher and based on student progress.**
 - The basic components of each lesson are reading and rereading easy or familiar books; analyzing student reading through running record techniques; composing, reading, and writing messages or stories; and reading a new book and demonstrating problem-solving strategies.

Reading Recovery®

- **Reading Recovery® has five outcome strategies:**
 1. discontinued students, who reach the average class literacy level and successfully leave the program;
 2. recommended action students, who receive the full program (60 lessons or 12 weeks), but are not successfully discharged from the program and may be recommended for another program;
 3. incomplete program students, who started the program at the end of the school year and are unable to complete it before year end;
 4. Moved students, who relocated before receiving the full program; and
 5. Other students, who do not fit into the preceding four categories.
- **In the analysis, we used “discontinued” students as the treatment group.**

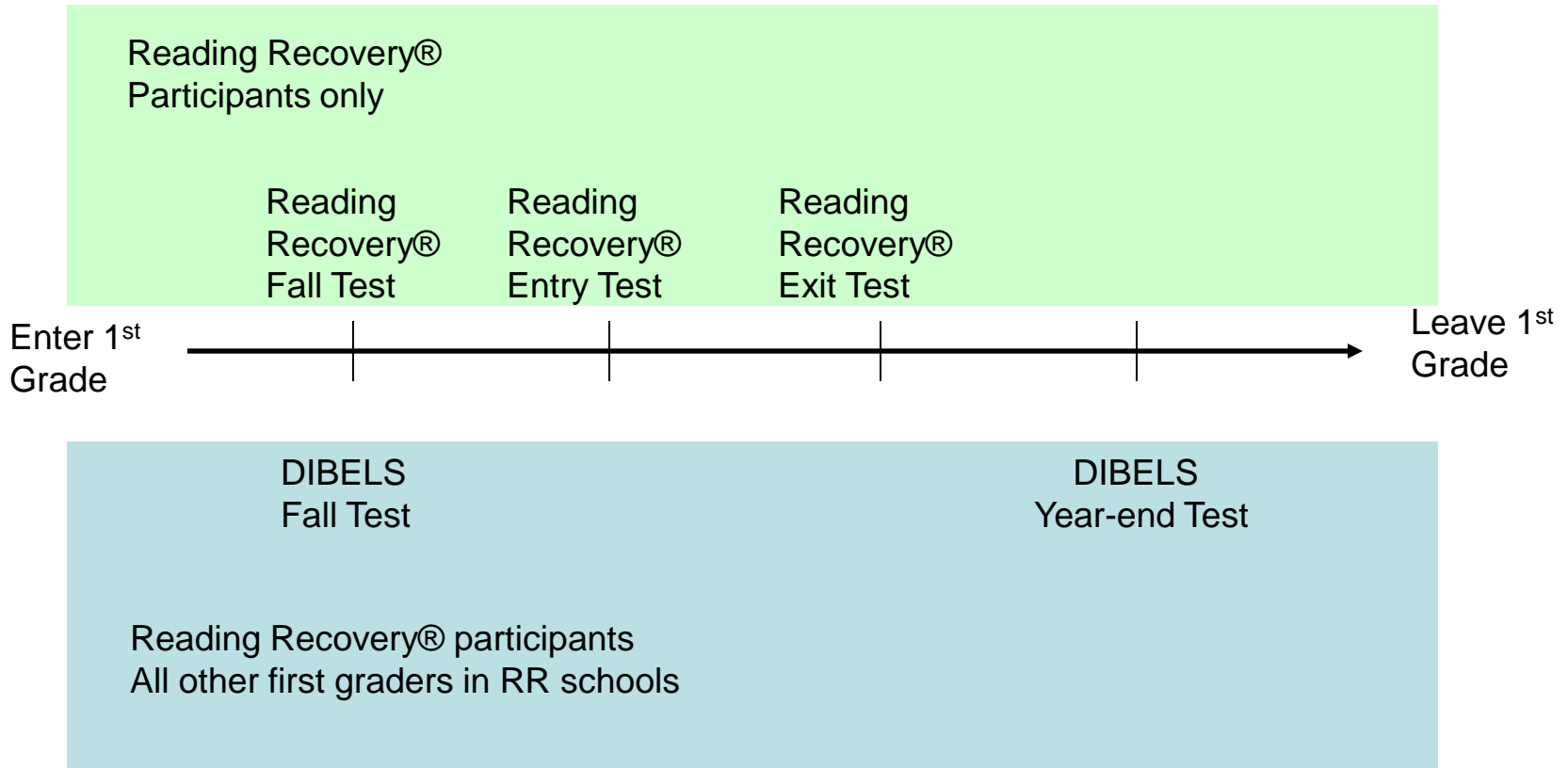
DIBELS

- **DIBELS (Dynamic Indicators of Basic Early Literacy Skills) tests:**
 - Phoneme Segmentation Fluency (PSF) assesses a student's ability to segment three- or four-phoneme words into their individual phonemes individually
 - Nonsense Word Fluency (NWF) measures ability to blend letters into words in which letters represent their common sounds
 - Word Use Fluency (WUF) measures the ability to use words correctly in a sentence and to understand the meaning of the word.
- **These tests are used as outcomes when we compare Reading Recovery® students with other first graders**
 - We cannot use Reading Recovery® tests because they are given only to Reading Recovery® students and consequently we have no comparable tests for non- Reading Recovery® students

Methodology

- **Observed outcomes, such as student test scores, are influenced by a variety of factors, including their participation in Reading Recovery®.**
- **Ideally, we would like to observe the expected outcome of the treated individuals were they not treated, but this is not observable.**
- **We need to find a comparison group that includes students who are as similar as possible to the students who participated in Reading Recovery®, yet did not participate in the program.**
- **Our approach is to construct comparison groups of first graders in KPS who have the same attributes (at least those we can measure) as the students who were selected to participate in Reading Recovery®.**
 - **These attributes include gender, race/ethnicity, free and reduced price lunch, special needs, and attended the same school as Reading Recovery® participants**
- **We then compare the outcomes, as measured by scores from various reading-related tests, of Reading Recovery® participants and non- Reading Recovery® students.**
- **We use first graders from the 2003/04 cohort and follow them through first grade and into the beginning of second grade.**

Sequence of Test Scores and Test Takers



Comparison Groups

Reading Recovery® Participants only

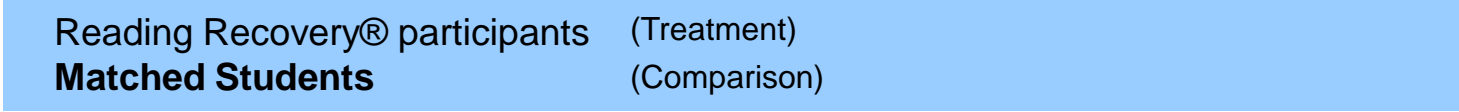
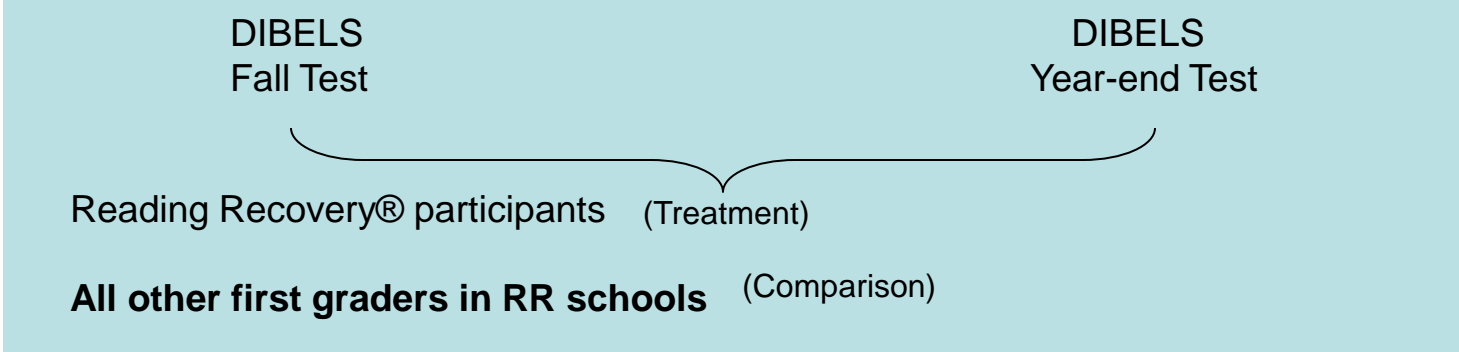
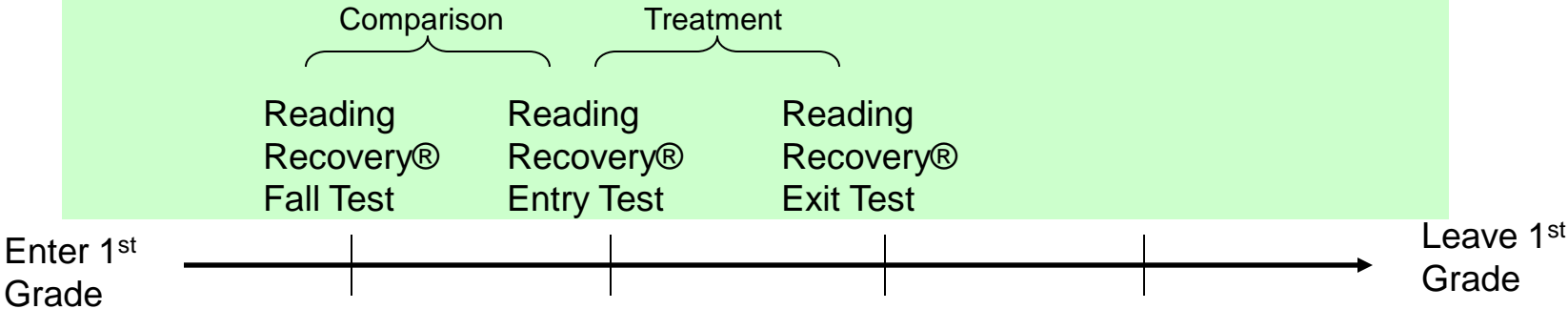
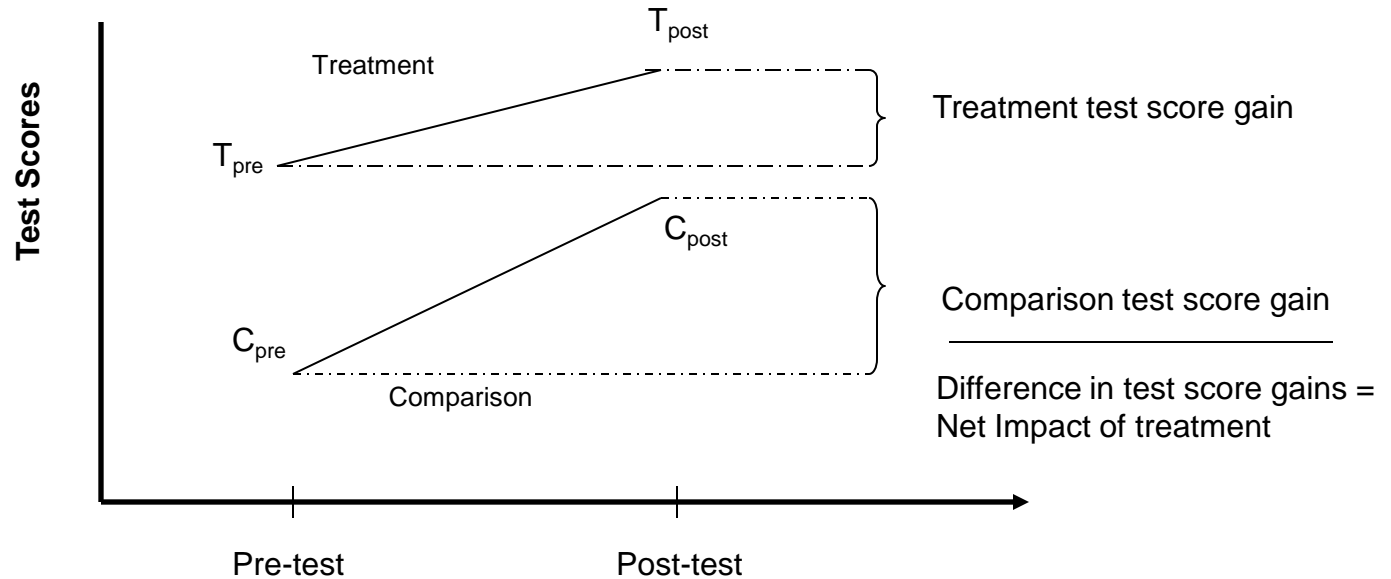
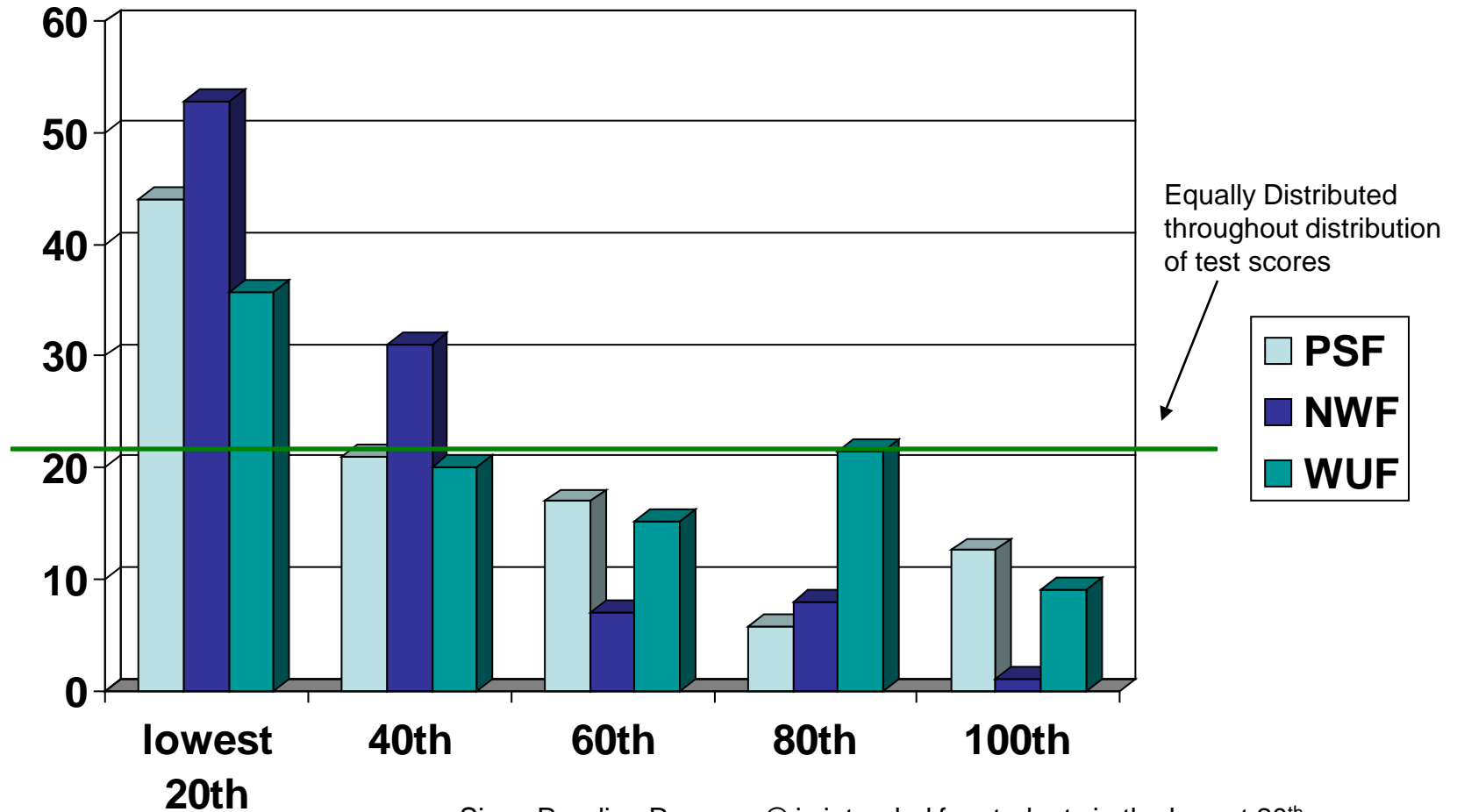


Illustration of Net Impact Analysis



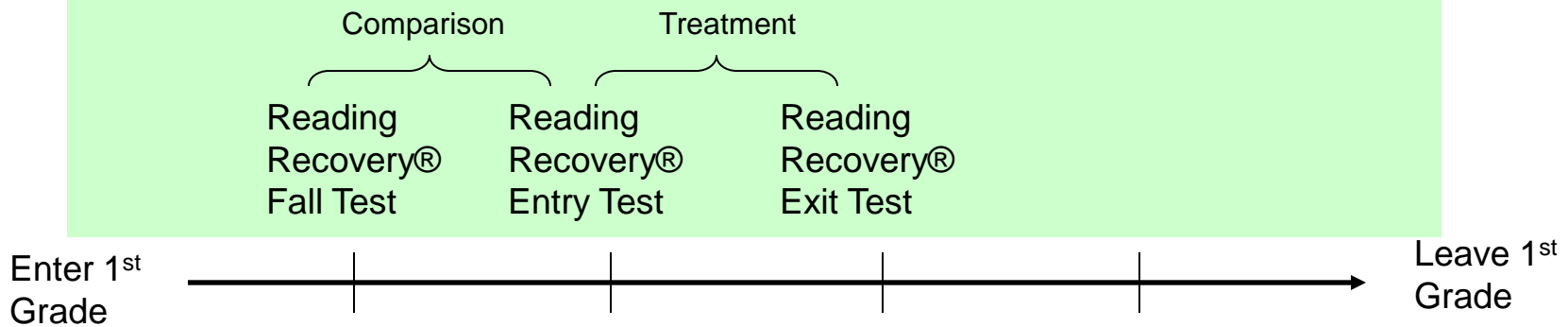
Distribution of Reading Recovery® Students by Beginning DIBELS Scores



Since Reading Recovery® is intended for students in the lowest 20th percentile of reading ability, we expect RR students to be much more highly represented in the lower quintiles than in the upper quintiles.

Comparison Groups: Reading Recovery® Students Only

Reading Recovery® Participants only

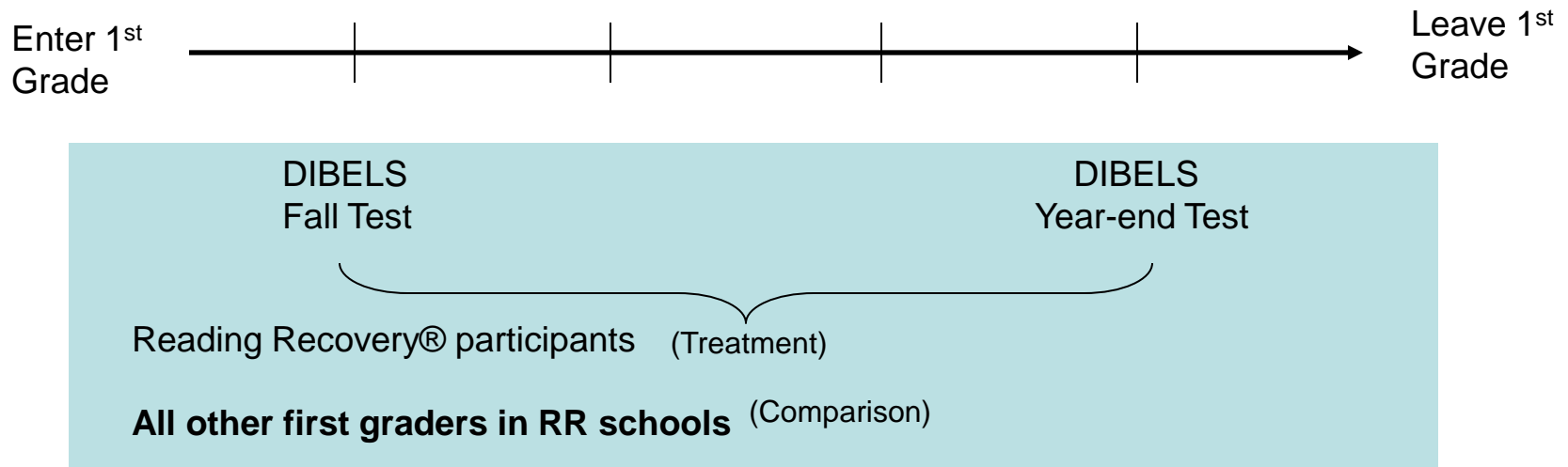


For this comparison group, we use the students who eventually were selected and enrolled in the Reading Recovery® program. A relatively large group of RR students did not enroll in the program until well into the school year. This offered an opportunity to compare their test score gains from the beginning of the school year until they enrolled (RR Entry Test-RR Fall Test) with their test score gains from the time they enrolled in RR until they successfully completed the program (RR Exit Test-RR Entry Test). Recognizing that the elapsed time within each episode may affect test score gains and that the initial score may affect gains, we divided the percentage gain in test scores by the number of days in each episode.

Reading Recovery® Students as Treatment and Control

	Difference in Level			(Difference in Level)/ Beginning level			(Difference in Level)/ Beginning Level/ Number of days between tests		
	Entry-1 st begin	Exit-Entry	T-stat	Entry-1 st begin	Exit-Entry	T-stat	Entry-1 st begin	Exit-Entry	T-stat
	Control	Treatment		Control	Treatment		Control	Treatment	
Letter	7.07	4.84	(0.92)	0.23	0.25	(0.09)	0.0014	0.0019	(0.27)
Word	6.40	7.96	(0.67)	4.64	2.33	(1.01)	0.0319	0.0241	(0.52)
Print	3.30	6.53	(2.54)	0.47	0.91	(1.19)	0.0032	0.0085	(1.62)
Vocabulary	13.7	14.56	(0.23)	3.42	1.90	(1.34)	0.022	0.016	(0.87)
Sound	12.7	9.1	(0.96)	3.15	0.80	(1.48)	0.020	0.008	(1.21)
Reading	2.9	9.9	(6.6)	2.2	3.0	(0.62)	0.014	0.030	(1.70)

Comparison Groups: All Other First Graders in RR Schools

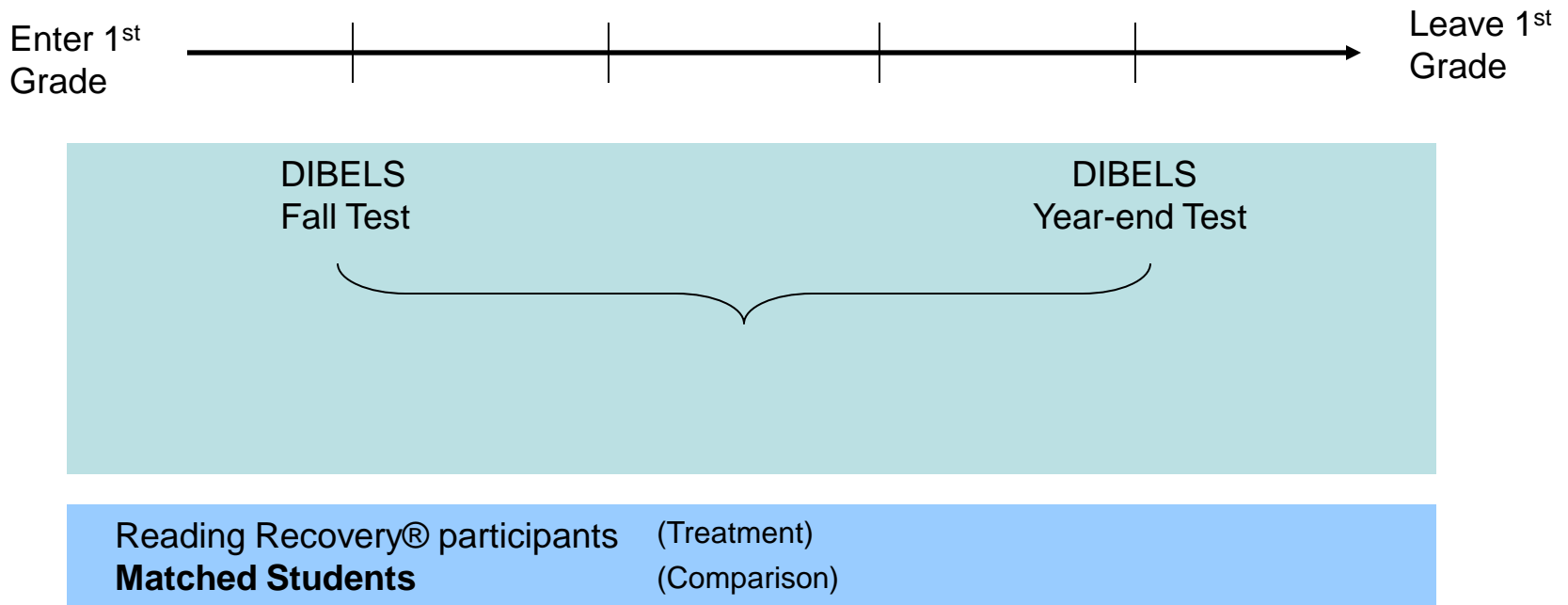


All Other First Graders as Comparison Group

DIBELS		Fall	1 st End	Difference	2 nd Beginning	Difference	Obs.
PSF	Treatment	21.7	55.7	34.0			33
	Comparison	32.7	51.3	18.6			700
	Difference	-11.0	4.4	15.4 (4.65)			
NWF	Treatment	8.4	61.9	53.5	51.8	43.4	33
	Comparison	24.5	61.1	36.6	55.5	31.0	700
	Difference	-16.1	0.8	16.9 (3.77)	-3.7	12.4 (2.01)	
WUF	Treatment	20.7	50.7	30.0			30
	Comparison	25.2	52.5	27.7			512
	Difference	4.5	1.8	2.3 (0.85)			

For two of the three DIBELS tests, the test score gains of RR students outpaced those of the comparison group, and for one that difference persisted into second grade. It is also interesting to note that RR students' scores approached or equaled those of the comparison group.

Comparison Groups: First Graders Matched with RR Students



Matched Comparison Group

- **Propensity Score Matching**

- The predicted probability of being in Reading Recovery was estimated for all the 86 Reading Recovery students and the rest of the first graders from the 10 schools where Reading Recovery was administered.
- The estimation is based on students' gender, race, lunch cost status, DIBELS test results at beginning of first grade, and the school building they attend.
- The nearest neighbor matching with replacement was used to select the comparison group—49 students were matched to 71 Reading Recovery students (15 of the 86 had missing test scores so were not used).
- To ensure the treatment group students are paired with students from their same building, the match was done for each building separately.
- The final matched data set includes 146 students, with 122 in the matched sample. Two of the original random sample students were selected as members of the comparison group.

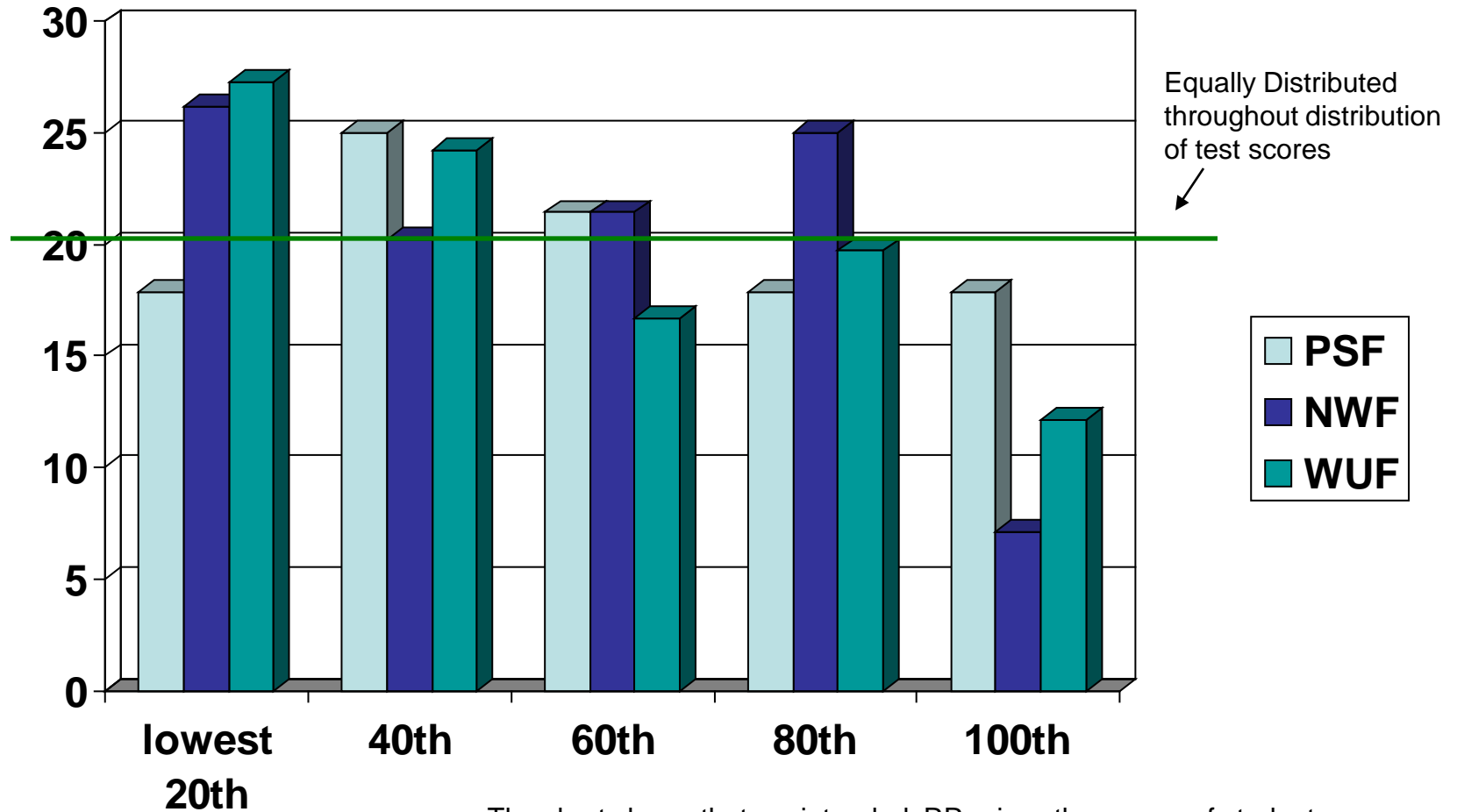
	All First Graders			Matched Sample		
	Treatment	Comparison	Difference	Treatment	Comparison	Difference
Male	0.63	0.50	0.13 (2.44)	0.62	0.65	-0.03 (0.31)
Black	0.56	0.42	0.14 (2.83)	0.50	0.63	-0.13 (1.25)
Hispanic	0.09	0.10	-0.01 (0.47)	0.12	0.06	0.06 (1.14)
Free lunch program	0.88	0.58	0.30 (5.99)	0.91	0.88	0.03 (0.57)

Matched First Graders as Comparison Group

DIBELS							
		Fall	1 st End	Difference	2 nd Beginning	Difference	Obs.
PSF	Treatment	21.7	55.7	34.0			33
	Comparison	22.7	48.0	25.3			86
	Difference	-1.0	7.7	8.7 (2.18)			
NWF	Treatment	8.4	61.9	53.5	51.8	43.4	33
	Comparison	8.1	45.6	37.5	40.0	31.9	73
	Difference	0.3	16.3	16.0 (4.14)	11.8	11.5 (2.20)	
WUF	Treatment	20.7	50.7	30.0			30
	Comparison	16.9	50.0	31.1			56
	Difference	4.5	1.8	-1.1 (0.60)			

The results for the matched students are similar to those for all other first graders. Also, not shown in this analysis, controlling for other characteristics such as gender, race, and free and reduced lunch status, did not change the results.

Distribution of Reading Recovery® Students by Ending DIBELS Scores



The chart shows that, as intended, RR raises the scores of students identified at the beginning of the school year with possible literacy deficiencies to levels similar to their peers by the end of the year.

Summary

- **The analysis estimates the net impact of Reading Recovery® as it was conducted in KPS for first graders in the 2003/04 cohort.**
- **Estimates show that students selected for Reading Recovery® are as much as twice as likely to be in the lowest 20th percentile of the distribution of reading-related test scores as students not selected for the program.**
- **At the end of the program, test scores of Reading Recovery® students are much more representative of the distribution of scores for the entire first grade class**
 - This result suggests that Reading Recovery® has helped students who demonstrated difficulties with literacy skills at the beginning of the year to perform at levels much closer to their peers by the end of the year.
- **Three comparison groups were used in the analysis: RR students only, all other first graders, and matched first graders.**
- **Analysis using all three comparison groups found evidence that students in Reading Recovery® raised their test scores more than students that were not in the program during the same period of time.**
 - These positive results were not found for all tests; for some of the tests the differences were statistically insignificant. In no case did we find a difference in favor of the comparison group that was statistically significant.
 - Controlling for other characteristics available in the dataset, such as free and reduced price lunch status, gender and ethnicity does not alter the results.
- **The analysis looked only at the net impact of the program and did not consider the cost effectiveness of Reading Recovery® nor did it compare the outcomes of Reading Recovery® to other reading programs.**