

IDEC Evaluation Report 2017–2018

Students Continue to Generate Strong Outcomes

Jeffrey B. Brymer-Bashore, The Ohio State University, International Data Evaluation Center

This report features the results of the 2017–2018 school year for the Reading Recovery® and Descubriendo la Lectura interventions in the United States. As described herein, Reading Recovery and Descubriendo la Lectura have continued to maintain strong outcomes, both in terms of progress across the length of the intervention, and as contrasted against comparison groups. These results are also comparable to those of the 2016–2017 school year (Lomax, 2018).

Summary of Reading Recovery Outcomes

Characteristics of participants

During the 2017–2018 school year, Reading Recovery was implemented by 16 university training centers responsible for overseeing the intervention in schools located in 42 states (as shown in Table 1). More than 35,000 children were selected to participate in the one-to-one intervention. There were 4,526 teachers trained in Reading Recovery, and they delivered the intervention with support from 266 teacher leaders in 211 training sites serving over 900 school districts. There were a total of 2,975 schools implementing Reading Recovery, and these schools were located in urban (27%), suburban (33.8%), and rural areas (39.2%).

Demographic information for the participating Reading Recovery students reveals that 54% were boys and 70% were eligible for free or reduced-price lunch. Children represented different ethnic backgrounds, including 56% White, 19% Hispanic, 17% African American, 2% Asian American, 1% Native American, and 5% representing either multiple races or other ethnic backgrounds.

The professional experiences of Reading Recovery teachers participating in the annual data collection

process and reported as means include a mean of 20.2 years of teaching experience and a mean of 8.4 years teaching Reading Recovery and/or Descubriendo la Lectura. On average, these teachers taught 7.5 Reading Recovery children during the current school year, and 42.1 children beyond their Reading Recovery load. Thus, they instructed a mean total of 49.7 children accounting for all teaching roles/assignments during this academic year.

Results

The assessment used in this analysis of outcomes for Reading Recovery was *An Observation Survey of Early Literacy Achievement* (Observation Survey) (Clay, 2013). This was administered to Reading Recovery students, a random sample of comparison students, and a sample of tested-not-instructed (TNI) students at fall, mid-year, and year-end. TNI students were those students considered for Reading Recovery, tested with the Observation Survey in the fall and again at mid-year, but not selected to receive Reading Recovery by the middle of the school year. They were also tested at year-end and comprised a second comparison group.

Of the students who received a complete series of Reading Recovery

Table 1. Participation in Reading Recovery in the United States, 2017–2018

Entity	n
University Training Centers	16
Teacher Training Sites	211
States	42
School Systems	964
School Buildings	2,975
Teacher Leaders	266
Teachers	4,526
Reading Recovery Students	35,579
Random Sample for RR	2,754
Tested-Not-Instructed for RR	5,493

NOTE: Some students in the Control Group of the random assignment study did not receive Reading Recovery. Their data are excluded from results in other tables in this report but included here.

lessons, end-of-program outcomes were as follows:

- 70% ($n = 18,852$) reached at least average levels of reading and writing achievement and their intervention programs were successfully discontinued.
- 30% ($n = 7,989$) made progress, but did not demonstrate proficiency at average levels of reading and writing. These students were recommended for consideration of additional interventions. Most notable were 5,412 students who were recommended for small-group literacy instruction or intervention other than special education, and 1,790 students who were recommended for literacy-related special education services.

Of the students with incomplete Reading Recovery interventions, outcomes were as follows:

- 18% ($n = 6,441$) were still in lessons at year-end without enough time in the school year to complete the intervention.
- 4% ($n = 1,303$) moved during the school year while still enrolled in lessons.

The data were further examined to explore two critical questions regarding the effectiveness of the Reading Recovery intervention. The first question is whether Reading Recovery students reach average levels of literacy achievement at the end of first grade as compared to all other first-grade children who did not receive the intervention. Here the average Observation Survey scores of Reading Recovery students were compared against all random sample students. (The Observation Survey is one met-

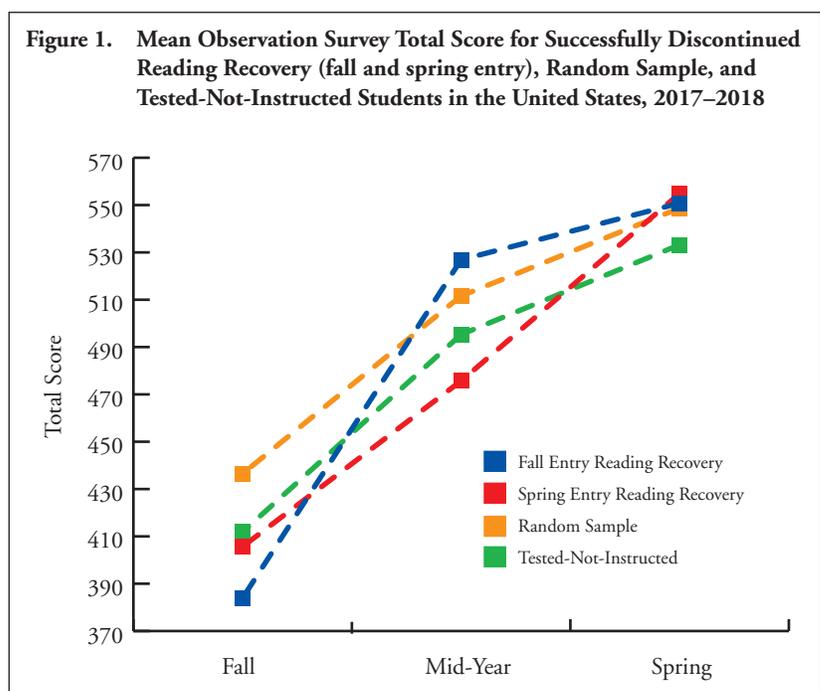
ric of literacy achievement level.) The second question is whether Reading Recovery students performed better at the end of the intervention than they would have performed if they were not provided the intervention. Here the average Observation Survey scores of Reading Recovery students were compared to the Observation Survey scores of TNI students.

These analyses were conducted using a Total Score derived from the Observation Survey. The Total Score scale of the Observation Survey was created based on 2009–2010 random sample student data (including the random sample students who received Reading Recovery). Students' Observation Survey scores on all six tasks from fall, mid-year, and year-end were used to create the total measure. The six tasks are Text Reading Level, Writing Vocabulary, Hearing and Recording Sounds in Words, Letter Identification, Ohio Word Test, and Concepts About Print. Instead of using the Observation Survey

scores of each student from the three time points, the random sample was divided into three randomly assigned groups, and the fall, mid-year, or year-end Observation Survey scores were chosen from each group, respectively, to represent an independent sample of students from the three time points during the school year.

The six tasks were treated as partial credit 'items' in a Rasch-based item response theory (IRT) analysis to convert the total raw scores to log-odd values ranging from approximately -4 to 4 . Those values were then converted through a linear transformation to create the final 0 to 800-point scale. As student scores were from various test points during the school year, the scale reflects yearlong growth. Thus, for example, a Total Score of 500 indicates the same literacy achievement level at any time point. Additional details on the Observation Survey (e.g., scale construction, reliability and validity evidence, normality,

Figure 1. Mean Observation Survey Total Score for Successfully Discontinued Reading Recovery (fall and spring entry), Random Sample, and Tested-Not-Instructed Students in the United States, 2017–2018



Strong effects such as these would not be possible without the strong commitment of our Reading Recovery and Descubriendo la Lectura trainers, teacher leaders, and teachers, who consistently seek to improve their teaching craft.

equal interval scales, unidimensionality) are described in D'Agostino (2012) and D'Agostino, Rodgers, and Mauck (2017).

Figure 1 shows the mean Observation Survey Total Scores for successfully discontinued Reading Recovery students served first (fall entry) during the school year, Reading Recovery students served second (spring entry), random sample students, and TNI students. Only students with valid scores at all three tests points were included in the analysis. As in past years, the TNI group had a slightly higher fall mean score relative to fall and spring entry Reading Recovery students, but not as high as the random sample students.

To examine the results, first consider the fall entry Reading Recovery students. By mid-year, these students had a greater mean gain than spring entry students, TNI students, and random sample students. Thus, the fall entry scores of Reading Recovery students, whose mean Observation Survey score was the lowest of all groups initially, were highest by mid-year. From mid-year to year-end, the average growth rate of the Reading Recovery fall entry students was slightly less than the average random sample student growth rate over the same period, but the two groups finished the year at about the same

achievement level, and both groups were considerably higher than TNI students.

Consider next the spring entry Reading Recovery students. These students had a smaller fall-to-mid-year mean gain than TNI students. Their lower performance was to be expected as this group did not receive the intervention until the second half of the school year. Thus, during the fall, the spring entry students serve as an additional control. Once they begin their intervention in the second half of the year, spring entry students had the largest growth rate. In addition,

the fall entry, spring entry, and random sample means were approximately the same at year-end testing, indicating that the Reading Recovery students had caught up to their random sample peers.

Figure 2 shows the results for the same four groups across the same three time points for Text Reading Level. The general trend as shown in Figure 2 is similar to that for the Observation Survey Total Score. The Reading Recovery discontinued students (both fall and spring entry) at year-end testing had not totally caught up to the random sample students. Note, however, that Reading Recovery students (both fall and spring entry) had reached grade-level expectations and had nearly achieved the text reading level of the random sample students.

Further analyses examined the means and magnitude of mean differences (effect sizes) at fall and year-end testing between the Reading Recovery students and the random sample or

Figure 2. Mean Text Level Score for Successfully Discontinued Reading Recovery (fall and spring entry), Random Sample, and Tested-Not-Instructed Students in the United States, 2017–2018

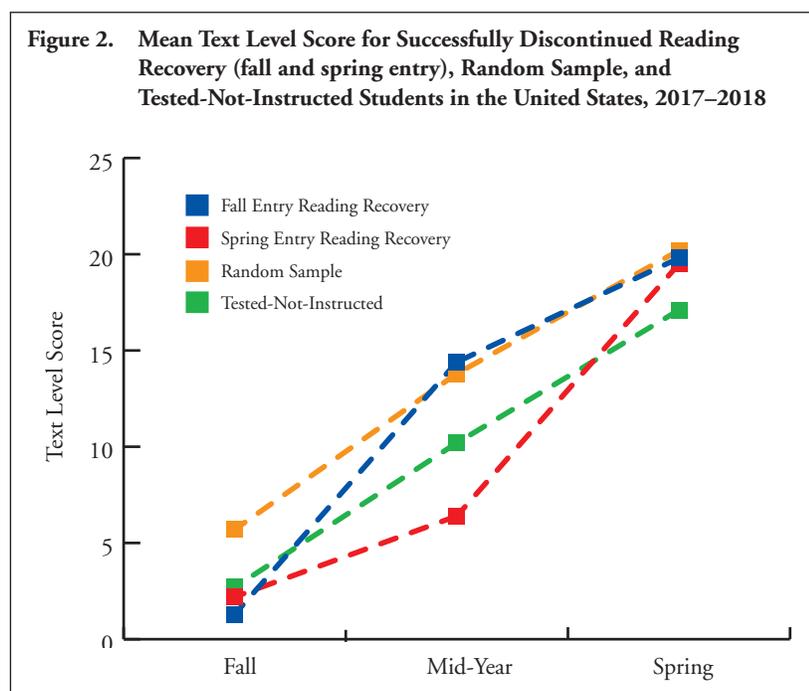


Table 2. Mean Fall and Year-End Total Scores with Effect Sizes for Successfully Discontinued Reading Recovery and Random Sample Students, 2017–2018

Observation Survey Task	Discontinued		Random Sample		Effect Size (<i>d</i>)	
	Fall	Year-End	Fall	Year-End	Fall	Year-End
Total Score	392.7	552.2	437.7	548.8	-0.82	0.07
Text Reading Level	1.7	19.7	5.8	20.2	-0.64	-0.06
Writing Vocabulary	12.3	55.6	20.6	54.8	-0.67	0.04
Hearing and Recording Sounds in Words	23.2	36.0	28.7	35.4	-0.67	0.17
Letter Identification	49.1	53.5	50.8	53.4	-0.33	0.07
Ohio Word Test	4.5	19.2	9.7	18.6	-0.83	0.15
Concepts About Print	13.0	21.0	15.1	20.5	-0.59	0.20

TNI students. Tables 2 and 3 display the total and individual task scores of fall entry and spring entry Reading Recovery discontinued students pooled together as compared with the random sample and TNI students respectively. For both tables, the far right-hand columns denote the effect sizes in terms of standardized mean differences. (Positive values indicate that the Reading Recovery mean was greater than the comparison mean value.) The effect size measure utilized was Cohen’s *d* (Cohen, 1988; Lomax & Hahs-Vaughn, 2012) which can be thought of in the metric of a standard deviation. Thus, a value of *d* = +1.00 would indicate that

the Reading Recovery children had a mean score of one standard deviation above the comparison group. A common standard to judge *d* is that .2 is a small effect size, .5 a medium effect size, and .8 a large effect size.

As displayed in Table 2, mean Reading Recovery students’ fall scores on all measures were substantially lower than the random sample, with medium to very large effect sizes (ranging from -.33 to -.83). By year-end testing, there were relatively small effect sizes in favor of the Reading Recovery students (ranging from .04 to .20), except for Text Reading Level (-.06). Thus, the Reading Recovery sample

began substantially below the random sample in the fall and by year-end had surpassed them on all but the Text Reading Level measure. More specifically, the effect size changes from fall to year-end were as follows: Total Score (0.89), Text Reading Level (0.58), Writing Vocabulary (0.71), Hearing and Recording Sounds in Words (0.84), Letter Identification (0.40), Ohio Word Test (1.02), and Concepts About Print (0.79). Thus, the Reading Recovery sample, as compared to the random sample, increased by three-quarters of a standard deviation unit from fall to year-end across the measures (an average effect size change of 0.75).

Table 3. Mean Fall and Year-End Total Scores with Effect Sizes for Successfully Discontinued Reading Recovery and Tested-Not-Instructed Students, 2017–2018

Observation Survey Task	Discontinued		Tested-Not-Instructed		Effect Size (<i>d</i>)	
	Fall	Year-End	Fall	Year-End	Fall	Year-End
Total Score	392.7	552.2	412.5	533.2	-0.50	0.44
Text Reading Level	1.7	19.7	2.7	17.1	-0.44	0.37
Writing Vocabulary	12.3	55.6	15.7	49.5	-0.40	0.36
Hearing and Recording Sounds in Words	23.2	36.0	26.3	34.9	-0.39	0.29
Letter Identification	49.1	53.5	50.1	53.1	-0.20	0.14
Ohio Word Test	4.5	19.2	6.6	18.2	-0.48	0.30
Concepts About Print	13.0	21.0	13.8	19.5	-0.26	0.55

The fall and year-end test scores for Reading Recovery discontinued students (fall and spring entry combined) and TNI students are shown in Table 3. In fall testing, the Reading Recovery sample Total Score mean and individual task means were all lower than the comparison TNI group's scores, with effect sizes ranging from -.20 (small) to -.50 (medium). By year-end testing, the Reading Recovery students had surpassed the TNI students on all measures, with effect sizes ranging from .14 (small) to .55 (medium). Thus, the Reading Recovery sample began in the fall substantially below the TNI sample and by year-end had surpassed them for all measures.

More specifically, the effect size changes from fall to year-end were as follows: Total Score (0.94), Text Reading Level (0.81), Writing Vocabulary (0.76), Hearing and Recording Sounds in Words (0.68), Letter Identification (0.34), Ohio Word Test (0.78), and Concepts About Print (0.81). Thus, the Reading Recovery sample, as compared to the TNI sample, increased by nearly three-quarters of a standard deviation unit from fall to year-end averaged across the measures (an average effect size change of 0.73).

In addition to these results, examination of the national data reveal the following outcomes of interest:

- First, on the Observation Survey Total Score, the discontinued students demonstrated acceleration from the 22th percentile in the fall to the 45th percentile at year-end.
- Second, in regard to classroom teachers' reports of their read-

ing group placements of Reading Recovery students, the discontinued students' placement in average or higher reading groups increased from 15% in the fall to 85% in these groups by year-end.

- Third, only 2% (N = 328) of all discontinued Reading Recovery students (N = 18,644) were referred to and placed in special education services following the intervention.

These are indications of the efficacy of the Reading Recovery intervention. At year end, discontinued students (a) have accelerated their literacy learning and have demonstrated performance within an average range on the Observation Survey Total Score, (b) have moved to the average, above average, or well above average reading groups, and (c) are not found to be referred for special education services in large numbers.

Summary of Descubriendo la Lectura Outcomes

Characteristics of participants

The Descubriendo la Lectura intervention, the reconstruction of Reading Recovery in Spanish, is designed for first graders who receive their initial literacy instruction in Spanish. Table 4 provides details about participation in Descubriendo la Lectura in the United States. For the 2017–2018 school year, 470 Descubriendo la Lectura children were instructed by 72 teachers. These Descubriendo la Lectura students attended 72 schools in 29 school districts located in eight states. These teachers were supported

by 29 teacher leaders. In addition, of all Descubriendo la Lectura students served, 56% were boys, 96% were Hispanic, 95% qualified for free or reduced-price lunch. The schools these students attended were located in urban (53.6%), suburban (39.3%), and rural areas (7.1%).

Trained teachers had a mean of 19.2 years of teaching experience and 7.4 years of Descubriendo la Lectura and/or Reading Recovery teaching experience. These teachers taught from 2 to 8 Descubriendo la Lectura children on a daily basis (mode = 4), while teaching a mean of 6.3 Descubriendo la Lectura children across the school year. On average, they also taught 35.4 children in their other teaching roles for a mean total of 41.7 children assigned to them during this academic year.

Table 4. Participation in Descubriendo la Lectura in the United States, 2017–2018

Entity	n
University Training Centers	3
Teacher Training Sites	25
States	8
School Systems	29
School Buildings	72
Teacher Leaders	29
Teachers	72
DLL Students	470
Random Sample for DLL	256
Tested-Not-Instructed for DLL	0

NOTE: Some students in the Control Group of the random assignment study did not receive Descubriendo la Lectura. Their data are excluded from results in other tables in this report but included here.

Figure 3. Mean Instrumento de Observación Total Score for Successfully Discontinued Descubriendo la Lectura (fall and spring entry), and Random Sample Students in the United States, 2017–2018

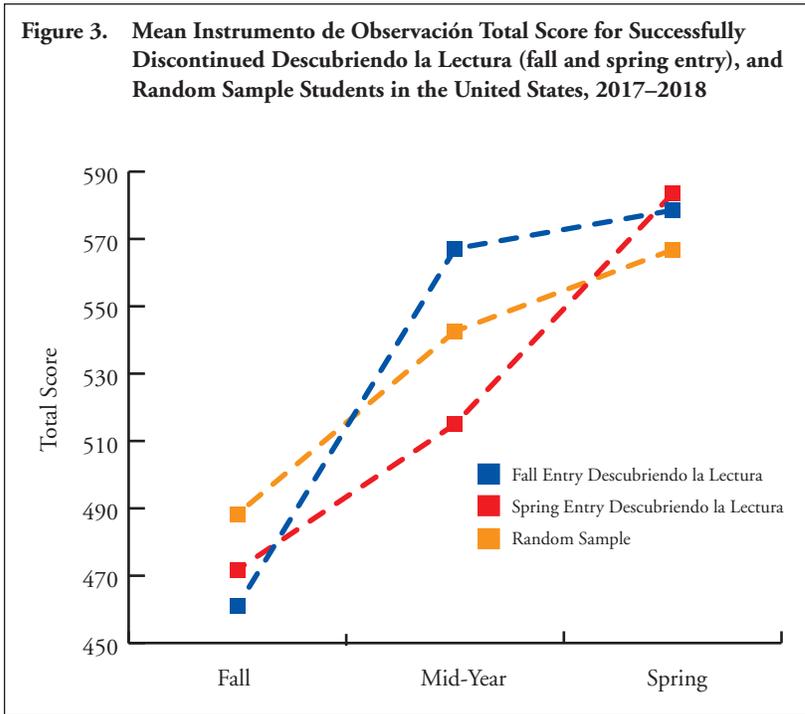
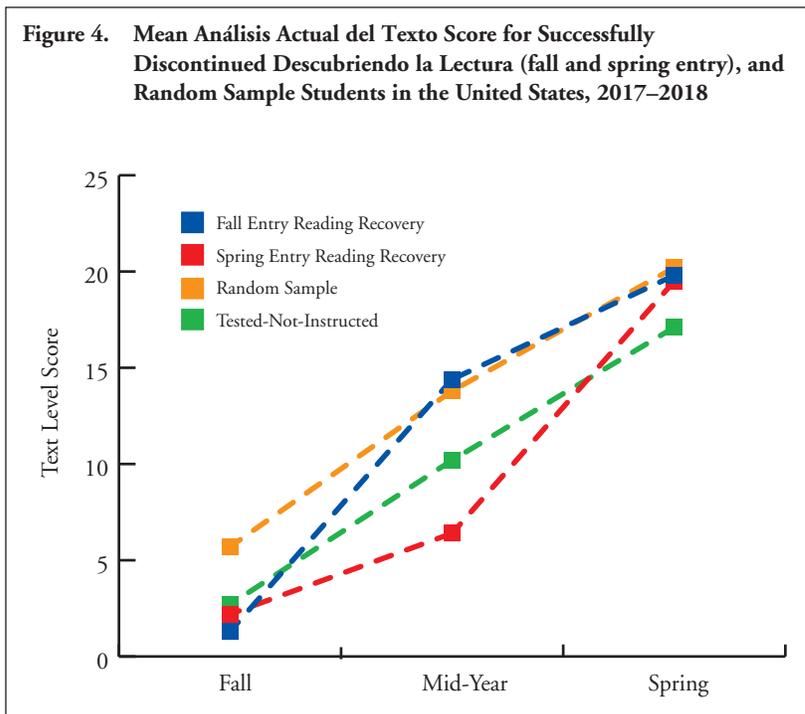


Figure 4. Mean Análisis Actual del Texto Score for Successfully Discontinued Descubriendo la Lectura (fall and spring entry), and Random Sample Students in the United States, 2017–2018



Results

The assessment used in this analysis of outcomes for Descubriendo la Lectura was *Instrumento de observación de los logros de la lecto-escritura inicial* (Instrumento de Observación) (Escamilla, et al., 1996). This was administered to both participating Descubriendo la Lectura students and a random sample of students for comparison purposes.

To secure a random sample, one half of the Descubriendo la Lectura schools were randomly chosen and two randomly selected students from each of these schools were administered the Instrumento de Observación. This random sample was the only comparison group available for the current analyses. Descubriendo la Lectura schools had last collected TNI data in 2011–2012, but due to very small samples in subsequent years leading to unstable average scores, IDEC has not continued ongoing, annual TNI testing and related data analyses.

Of all students served in Descubriendo la Lectura, 41% ($n = 193$) reached the average reading levels of their peers and thus were discontinued successfully. Another 27% ($n = 125$) were recommended for further evaluation, 2% ($n = 11$) moved, and 27% ($n = 125$) received incomplete interventions. Of the students who completed the intervention (both discontinued and referred students), 61% were discontinued. In regard to the referred students, of note were 69% recommended for small-group literacy instruction or intervention other than special education. Sixteen percent were recommended for literacy-related special education services.

Table 5. Mean Fall and Year-End Total Scores with Effect Sizes for Successfully Discontinued Descubriendo la Lectura and Random Sample Students, 2017–2018

Instrumento de Observación Task	Discontinued		Random Sample		Effect Size (<i>d</i>)	
	Fall	Year-End	Fall	Year-End	Fall	Year-End
Total Score	465.1	580.2	488.3	566.8	-0.50	0.36
Análisis Actual del Texto	1.0	19.1	4.0	17.5	-0.62	0.21
Escritura de Vocabulario	10.3	48.3	17.5	44.0	-0.60	0.25
Oír y Anotar los Sonidos en las Palabras	24.4	38.4	29.1	37.3	-0.43	0.30
Identificación de Letras	46.2	59.2	49.0	57.5	-0.22	0.30
Prueba de Palabras	6.9	19.6	11.2	18.3	-0.62	0.31
Conceptos del Texto Impreso	11.4	20.5	12.9	19.6	-0.36	0.26

For further analyses, the random sample students' scores on the six tasks of the Instrumento de Observación were combined to create a Total Score (with a 0 to 800-point range) that reflects literacy development throughout the school year. This parallels the processes applied to Reading Recovery data described earlier.

Among the fall entry, spring entry, and random sample groups, the largest growth from fall to mid-year on the Instrumento de Observación Total Score was for the fall entry Descubriendo la Lectura students (see Figure 3). From mid-year to year-end, the largest growth was for the spring entry Descubriendo la Lectura students. Together these results indicate that the greatest gain for all students observed was during the respective Descubriendo la Lectura intervention periods. Spring entry students and random sample students showed approximately the same gain from fall to mid-year. However, from mid-year to year-end, the spring entry Descubriendo la Lectura students outgained the random sample.

The trend for Text Level scores (see Figure 4) was very similar to the Total Score trend. By year-end testing, both fall and spring entry Descubriendo la Lectura students had substantially surpassed the scores on both measures as compared to the random sample group. In other words, both Descubriendo la Lectura groups began the school year behind the random sample, but caught up to and exceeded the random sample group by the end of the year.

In Table 5 are the mean scores and effect sizes (Cohen's *d*) for fall and spring entry Descubriendo la Lectura discontinued students combined, as well as the random sample students at both fall and end of year testing. In fall testing, the Descubriendo la Lectura sample, Instrumento de Observación Total Score mean, and individual task means were all lower than the comparison random sample group, with effect sizes ranging from $-.24$ (small) to $-.65$ (medium). By year-end testing, the Descubriendo la Lectura students had surpassed the random sample students on all

measures, with effect sizes ranging from $.21$ (small) to $.35$ (small). Thus, the Descubriendo la Lectura sample began the fall substantially below the random sample and by year-end had surpassed them on all measures.

More specifically, the effect size changes for the Descubriendo la Lectura students and random sample students from fall to year-end were as follows: Instrumento de Observación Total Score (.86), Text Reading Level (Análisis Actual del Texto, .83), Writing Vocabulary (Escritura de Vocabulario, .85), Hearing and Recording Sounds in Words (Oír y Anotar los Sonidos en las Palabras, 0.73), Letter Identification (Identificación de Letras, 0.52), Ohio Word Test (Prueba de Palabras, 0.93), and Concepts About Print (Conceptos del Texto Impreso, 0.62). Overall, the Descubriendo la Lectura sample, as compared to the random sample, increased by slightly more than three-quarters of a standard deviation unit from fall to year-end averaged across the measures (an average effect size change of 0.76).

Other results noted in the data include the following:

- First, on the Instrumento de Observación Total Score the discontinued students have accelerated their learning as shown in their progression from the 24th percentile in the fall to the 53rd percentile at year-end.
- Second, when considering the classroom reading group placements assigned by their teachers, the discontinued students' placements increased from 12% in the average or higher reading groups in the fall to 92% in such groups by year-end.
- Finally, only 1% (N = 1) of all discontinued students (N = 189) were referred and placed in special education following the intervention.

These are additional indications of the efficacy of the Descubriendo la Lectura intervention, as discontinued students (a) have accelerated their literacy learning and have achieved an average Total Score at year-end, (b) have been advanced to the average, above average or well above average reading groups, and (c) are not referred for special education services in large numbers.

Conclusion

The results reported here for the Reading Recovery and Descubriendo la Lectura interventions, as well as prior results (e.g., Lomax, 2018), indicate that Reading Recovery and Descubriendo la Lectura continue to be amongst a very small list of educational interventions with strong impacts on student learning in the United States. Now in its 34th year of implementation in 2018–2019, students receiving these interventions continue to generate strong outcomes.

Strong effects such as these would not be possible without the strong commitment of our Reading Recovery and Descubriendo la Lectura trainers, teacher leaders, and teachers, who consistently seek to improve their teaching craft. The efforts of these educators continue to result in outstanding literacy success for participating students.

References

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Clay, M. M. (2013). *An observation survey of early literacy achievement* (3rd ed.). Portsmouth, NH: Heinemann.
- D'Agostino, J. V. (2012). Technical review committee confirms highest NCRTI ratings for Observation Survey of Early Literacy Achievement. *The Journal of Reading Recovery, 11*(2), 60–64.

About the Author

Jeff Brymer-Bashore is the director and co-principal investigator at the International Data Evaluation Center (IDEC) and has worked at IDEC for more than 18 years. With a degree in mathematics from The Ohio State University, he specializes in designing and managing large-scale data collections systems. When he is not working with data, Jeff is a DJ and children's entertainer.



- D'Agostino, J. V., Rodgers, E., & Mauck, S. (2017). Addressing inadequacies of the Observation Survey of Early Literacy Achievement. *Reading Research Quarterly*, doi: 10.1002/rrq.181
- Escamilla, K., Andrade, A. M., Basurto, A. G. M., Ruiz, O. A., & Clay, M. M. (1996). *Instrumento de observación de los logros de la lecto-escritura inicial*. Portsmouth, NH: Heinemann.
- Lomax, R. G. (2018). Results again show strong impact on student learning. *The Journal of Reading Recovery, 17*(2), 51–58.
- Lomax, R. G., & Hahs-Vaughn, D. L. (2012). *An introduction to statistical concepts* (3rd ed.). New York: Routledge.