The Albuquerque Reads Program: A Three-Site Analysis 2007-2009

Nancy Carrillo, PhD
May 2009

ALBUQUERQUE PUBLIC SCHOOLS

BOARD OF EDUCATION

MARTIN ESQUIVEL
President

PAULA MAES
Vice President

DOLORES GRIEGO
Secretary

DAVID EUGENE PEERCY
Instructional Policy Chair

LORENZO L. GARCIA
District Relations Chair

DAVID L. ROBBINS
Finance/Audit Chair

ROBERT D. LUCERO
Capital Outlay Chair

Superintendent
WINSTON BROOKS

LINDA SINK
Chief Academic Officer

BRAD WINTER
Chief Operations Officer

RUBY ETHRIDGE
Associate Superintendent

DIANE KERSCHEN
Associate Superintendent

RAQUEL REEDY
Associate Superintendent

EDUARDO SOTO
Associate Superintendent

INSTRUCTION AND ACCOUNTABILITY

Rose-Ann McKernan, Executive Director

Research, Deployment & Accountability

Thomas Genné, Director

6400 Uptown Blvd. NE (400 EAST)
Albuquerque, New Mexico 87110
(505) 872-6870
www.rda.aps.edu

EVALUATION REPORT

NANCY CARRILLO *MAY 2009*

THE ALBUQUERQUE READS PROGRAM: A THREE-SITE ANALYSIS 2007-2009

Introduction

The Albuquerque Reads program is a partnership between the Greater Albuquerque Chamber of Commerce (GACC) and Albuquerque Public Schools (APS) designed to reinforce reading skills among all kindergarten students. The program consists of 30 minutes of structured one-on-one tutoring by community volunteers and 30 minutes of teacher-led, small group instruction each week.[1] All kindergarten students in three schools, Atrisco, Bel-Air, and Wherry Elementaries, participate. Previous research suggests that for the 2006-07 SY, Albuquerque Reads' students gained about five more points on the Kindergarten Developmental Progress Report (KDPR) from fall to spring after controlling for demographic factors. The cost of the program approximated one program similarly designed for kindergarten students at about $1,200 per student. For more details about the program and outcome results please read the previously published evaluation (Carrillo, 2008).[2]

Research Question

1. Does the program 'Albuquerque Reads' improve reading ability among kindergarteners?

We assessed this question for both the winter and spring assessment windows. As shall be revealed, the analysis of KDPR Reading scores led to another research question and a deeper analysis. To foreshadow, we see a substantial impact of the Albuquerque Reads program on spring reading scores, yet no impact on winter scores. We explore three possible explanations for this curious finding.

Method

We compared reading assessment scores of Albuquerque Reads and comparison (or 'peer') students for 2007-08 and 2008-09 kindergarten students. The 'Albuquerque Reads' group consisted of all kindergarten students at Atrisco, Wherry, and Bel-Air Elementaries. The comparison group consisted of all kindergarten students in all other schools in the same Statistical Peer groups, which were Peer groups 1, 2, and 4.[3] For 2007-08 students, we collected

[1] In some cases tutoring may be two-on-one due to a lack of volunteers.
[2] Carrillo, N. 2008. "Albuquerque Reads: A Three-Site Analysis," Albuquerque Public Schools. http://www.rda.aps.edu/RDA/Documents/Publications/07_08/Albuquerque_Reads_2006_07.pdf.
[3] Statistical Peers for Benchmarking groups schools with similar student compositions based on each school's percentage of students participating in the Free and Reduced Price Meals program, percentage of students who are English Language Learners, and percentage of students who are a member of an underperforming minority ethnic group. The most recent analysis places elementary schools into seven groups of between seven and seventeen

Fall '07, Winter '08, and Spring '08 KDPR Reading scores from Albuquerque Reads and comparison schools. For 2008-09 kindergarten students, we collected Fall '08 and Winter '09 assessment scores. We calculated total reading score as well as proficiency level.[4] Analyses involving the winter assessment were limited to students who remained in the same school for both the fall and winter assessments; while analyses involving the spring assessment were limited to students who remained in the same school throughout the entire year. For students assessed in both English and Spanish, the lower scores were omitted.

Results

A preliminary way to assess the impact of Albuquerque Reads is to examine KDPR Reading proficiency levels. Table 1 shows the percentage of Albuquerque Reads and comparison students who are proficient in each available window. For both cohort groups and languages, the percentages of students who are proficient are close to zero in the fall. In the winter, the percentage of Albuquerque Reads students who are proficient tends be *lower* than the percentage of comparison group students. By the spring, proficiency levels are somewhere around 50%; Albuquerque Reads and comparison groups do not differ.

Table 1: KDPR Reading Proficiency for Albuquerque Reads and Comparison Group Students

		2007-08 Cohort				**2008-09 Cohort**		
Language	Treatment	*n*	Fall 2007	Winter 2008	Spring 2008	*N*	Fall 2008	Winter 2009
English	Albuquerque Reads	178-193	0%	17%	57%	152-164	5%	34%
	Comparison Group	2,101-2,167	1%	22%	57%	1,948-2,048	2%	34%
Spanish	Albuquerque Reads	14-25	0%	4%	48%	11	0%	0%
	Comparison Group	540-575	1%	12%	48%	514-562	0%	19%

Table 1 loses information by considering proficiency levels, a measure of only two values - proficient or not, rather than raw scores that range from 0 to either 24 (in the fall) or 33 (in the winter and spring). Thus, while we see how many students become proficient over time, we cannot detect how much change has occurred, nor do we know whether the groups were approximately academically equivalent at the start of their kindergarten year. In fact,

---

schools. Dunavin, R. (2008). "Statistical Peers for Benchmarking 2008," Albuquerque Public Schools. http://www.rda.aps.edu/RDA/Documents/Publications/08_09/Statistical_Peers_for_Benchmarking_2008.pdf

[4] The KDPR reading assessment changed from 2007-08 to 2008-09. Cut scores changed in the fall, winter, and spring; and the items assessed in the winter session also changed. For the purposes of this report, for both years, the fall reading scores are comprised of KDPR questions 33, 34, 37, 40, 42, 43, and 44. Fall cut scores are: Area of Need (less than 16), Emergent (16-18), Not Proficient (19-21), and Proficient (above 21). The winter reading scores are comprised of questions 33, 34, 35, 36, 37, 38, 39, 40, 42, 43, and 44. Winter cut scores are: Area of Need (less than 20), Emergent (20-24), Not Proficient (25-27), and Proficient (above 27). The spring reading scores are comprised of questions 33, 35, 36, 37, 38, 39, 40, 41, 42, 43, and 44. Spring cut scores are: Area of Need (less than 19), Emergent (19-24), Not Proficient (25-27), and Proficient (above 27).

Albuquerque Reads students appear to be behind their peers from the outset, a fact hidden in Table 1 since almost no students in either group are proficient in the fall. The fall KDPR Reading raw scores in Table 2 indicates that peers score between 1.5 and 3.9 points higher than Albuquerque Reads students, substantial and usually statistically significant differences.

Table 2: Average Fall Raw KDPR Reading Points

Language	Treatment	Fall 2007	Fall 2008
English	Albuquerque Reads	4.9 (*n*=193)	8.2 (*n*=164)
	Comparison Group	8.8 (*n*=2,101)	9.7 (*n*=2,048)
	Statistically significant difference?	*t*=11.8*	*t*=3.0*
Spanish	Albuquerque Reads	2.9 (*n*=14)	5.5 (*n*=11)
	Comparison Group	5.5 (*n*=575)	7.3 (*n*=560)
	Statistically significant difference?	*t*=6.8*	*t*=1.2

* $p<.05$

Table 3 provides a more meaningful growth analysis be examining the change in KDPR Reading scores over time for Albuquerque Reads and peer students. Since the number of possible points students can earn changes with the testing window (i.e. fall, winter or spring), we examine the gain in the percentage of points students earn. The table shows the average gain in KDPR Reading percentage points for Albuquerque Reads and comparison group students in English and Spanish.[5] Combining all students in the 2007-08 cohort, for instance, kindergarteners earned 33% of the possible points in the fall, 59% of possible points in the winter, and 80% of possible points in the spring, therefore gaining 26 percentage points in the winter and 47 percentage points in the spring compared to the fall.

Table 3: Average Gain in KDPR Winter and Spring Reading Percentage Points from Fall Percentage Points

		2007-08 Cohort		**2008-09 Cohort**
Language	Treatment	Gain in KDPR, Winter	Gain in KDPR, Spring	Gain in KDPR, Winter
English	Albuquerque Reads	38 (*n*=179)	60 (*n*=177)	30 (*n*=152)
	Comparison Group	21 (*n*=2,101)	43 (*n*=2,122)	26 (*n*=1,948)
Statistically significant difference?		*t*=7.2*	*t*=10.9*	*t*=1.7
Spanish	Albuquerque Reads	35 (*n*=27)	67 (*n*=25)	35 (*n*=11)
	Comparison Group	29 (*n*=574)	56 (*n*=540)	29 (*n*=514)
Statistically significant difference?		*t*=1.6	*t*=2.3*	*t*=.74

* $p<.05$

[5] For the purposes of this report, language is determined by the language of the spring administration of KDPR.

Looking first at students assessed in English, Albuquerque Reads students gained more percentage points than their peers in both the winter and spring windows of the 2007-08 school year as well as the winter window of the 2008-09 school year. Combining both cohorts, in the winter, Albuquerque Reads students gained on average about 34 percentage points, compared to about 24 percentage points among their peers. In the spring of 2007-08, Albuquerque Reads students gained 60 percentage points, compared to 43 among their peers. The differences are statistically significant and considerable for both windows of 2007-08. In 2008-09, the difference of 4 percentage points in the winter window is not statistically significant.

In Spanish, again Albuquerque Reads students gained more percentage points than their peers in all three windows. Differences tend to be smaller in magnitude, but one is substantial – the statistically significant difference of 11 percentage points in the spring of 2008.

Table 4 breaks the spring 2008 information down by school. In English, students in one of the three Albuquerque Reads schools earned more KDPR Reading percentage points on average from fall to spring than any other school in its statistical peer group. In the other two Albuquerque Reads schools, students earned the third highest average gain among their peers. Two of the Albuquerque Reads schools assessed ten or more students in Spanish. One of these earned more percentage points than all other schools in the statistical peer group. The other had the fourth highest gain among eight schools.

In sum, Albuquerque Reads students and their peers do not appear to be different when it comes to achieving proficiency by the end of the kindergarten year. But this finding belies the fact that students in Albuquerque Reads schools tend to start their kindergarten year academically under-prepared compared to their peers. Albuquerque Reads students show significantly more growth over the school year.

Table 4: Average Gain in KDPR Reading Percentage Points from Fall to Spring by School, 2007-08 Cohort

	English		Spanish	
School	**Avg. gain**	**# students**	**Avg. gain**	**# students**
Statistical Peer Group 1				
Albuquerque Reads	***67***	***78***	***56***	***10***
Comparison School 1	63	19	73	58
Comparison School 2	50	20	47	35
Comparison School 3	49	32	67	84
Comparison School 4	48	64	57	11
Comparison School 5	44	106	*n/a*	
Comparison School 6	42	83	56	24
Comparison School 7	41	36	47	30
Comparison School 8	38	39	30	16
Statistical Peer Group 2				
Comparison School 9	66	124	71	12
Comparison School 10	63	59	66	34
Albuquerque Reads	***53***	***39***	***74***	***15***
Comparison School 11	51	76	64	15
Comparison School 12	50	52	71	39
Comparison School 13	48	43	28	15
Comparison School 14	48	79	*n/a*	
Comparison School 15	47	62	*n/a*	
Comparison School 16	46	46	*n/a*	
Comparison School 17	45	14	*n/a*	
Comparison School 18	41	86	*n/a*	
Comparison School 19	37	63	37	28
Comparison School 20	35	70	52	18
Comparison School 21	31	17	53	19
Comparison School 22	30	166	57	40
Comparison School 23	29	44	35	29
Comparison School 24	25	40	*n/a*	
Statistical Peer Group 4				
Comparison School 25	67	76	*n/a*	
Comparison School 26	59	50	*n/a*	
Albuquerque Reads	***55***	***60***	***n/a***	
Comparison School 27	44	70	51	16
Comparison School 28	39	45	*n/a*	
Comparison School 29	39	56	*n/a*	
Comparison School 30	27	108	*n/a*	
Comparison School 31	26	81	*n/a*	
Comparison School 32	19	56	*n/a*	

Data for schools assessing fewer than 10 students in Spanish are omitted.

However, such a finding could be misleading. First, the fact that Albuquerque Reads students show more growth may simply be due to the fact that they started lower.[6] In addition, many other factors influence assessment scores, and one or more of these may be the real cause behind the positive Albuquerque Reads results. A more robust statistical test will include such factors. In Table 5, we use OLS regression to predict winter and spring KPDR Reading raw scores, controlling for students' previous achievement (in the form of the fall KDPR Reading raw score) as well as the important demographic variables gender, ethnicity, poverty status, status as a non-English speaker, and special education status. We predict girls, Anglos and Asians, non-poor students, non-ELL students, and non-special education students will score higher. While it is not expected that cohort, Statistical Peer group, or language of assessment will necessarily impact scores, we have controlled for these variables as well.[7] Clustering by location helps to control for variations between schools that are not accounted for in the model.

These analyses suggest that students in Albuquerque Reads schools gain no more points on the KDPR Reading winter assessment compared to students in other schools. Control factors influence winter scores as expected. Winter reading scores are strongly positively related to fall reading scores. Girls outperform boys slightly. Both ELL and special education status decrease scores. Students assessed in Spanish score higher than those assessed in English.[8]

On the other hand, Albuquerque Reads shows a strong positive impact on spring KDPR Reading scores, increasing them by about two and a half points. Spring scores are also higher among girls, Caucasians, and Asians; and scores are positively related to fall reading scores. Scores are again negatively related to ELL and special education status.

[6] Regression to the mean is a common program evaluation problem when, as is often the case, participants are compared to an *a priori* higher-achieving peer group. "Random variance affects the measurement of any variable: this random variance will cause some samples to be extreme. For the second measurement, these samples will appear to regress because the random variance affecting the samples in the second measurement is independent of the random variance affecting the first. Thus, regression toward the mean is a mathematical inevitability: any measurement of any variable that is affected by random variance *must* show regression to the mean." Wikipedia: http://en.wikipedia.org/wiki/Regression_to_the_mean, 27 April 2009 (original emphasis).

[7] Gender is a dummy variable where female is coded 1. Well-performing ethnicity is a dummy variable where underperforming minorities (Blacks, Hispanics and Native Americans) are coded 0, Asians and Whites are coded 1. Poverty status is a dummy variable in which students participating in Free or Reduced Lunch Program are coded 1. Special education status is a dummy variable; students receiving special education services are coded 1. Statistical Peer Groups 1 and 4 are dummy variables and are compared to students in Statistical Peer Group 2. Students assessed in Spanish are coded 1 in the dummy variable 'Assessed in Spanish'; and students in the second cohort are coded 1 in the dummy variable '2008-09 Cohort'.

[8] This finding may be an artifact. In this data set, a higher percentage of students with Spanish scores than those with English scores were, in fact, assessed in both languages. For such students, only their higher scores are represented in the data set. These students therefore have the advantage of having taken the test twice and having their higher scores included.

Table 5: Impact of Albuquerque Reads on Gains in KDPR, 2007-08

	Winter KDPR Reading Score	Spring KDPR Reading Score
Cohorts	2007-08 & 2008-09	2007-08 only
n	4418	2862
r^2	.35	.26
ABQ Reads	**-.12 (1.00)**	**2.45* (.65)**
Fall KDPR Reading Score	.68* (.06)	.58* (.06)
Gender (girls)	.81* (.19)	.83* (.29)
Well-performing ethnicity (Caucasians and Asians)	.35 (.21)	.63* (.26)
Poverty Status	-.65 (.87)	-.60 (.50)
Special Education Status	-3.30* (.55)	-3.39* (.57)
English Language Learners	-1.82* (.37)	-1.21* (.29)
Assessed in Spanish	1.89* (.66)	1.74 (.88)
Statistical Peer Group 1	.83 (.97)	.26 (.81)
Statistical Peer Group 4	.32 (1.44)	-.79 (1.11)
2008-09 Cohort	1.75 (1.23)	*n/a*

* $p<.05$. Standard errors in parentheses. Clustered by school.

The unexpected finding that Albuquerque Reads outcomes differ dramatically for winter and spring leads us to a second research question:

Research Question

2. What might account for the fact that the Albuquerque Reads program appears to influence spring scores substantially without influencing winter scores?

We discussed the above findings with district experts. They expressed three conflicting, though not necessarily mutually exclusive, hypotheses:

> *H1: Albuquerque Reads teachers are particularly pressured to score students high on the DRA2 assessment, whose results are included in the spring KDPR Reading score only.*

This hypothesis traces the difference between winter and spring results to the key difference in the KDPR Reading assessment itself. The 'reading at grade level' item, arguably the item of

most importance, is assessed with the DRA2 assessment. In the 2007-08 SY, this particular item was not measured in the winter window.[9] This hypothesis implies that Albuquerque Reads is not necessarily effective. In this view, positive findings are an artifact of biased test administration.[10] Biased test administration can occur intentionally or unintentionally, and may be a result of training differences.

> *H2: Albuquerque Reads improves reading and fluency, but does not necessarily improve other aspects of literacy included in KDPR.*

District literacy experts note that new readers, especially, may be confused by some pre-reading assessment tasks that comprise several of the other KDPR Reading assessment items, such as phoneme segmentation and nonsense word fluency, even if they had previously mastered these skills. This hypothesis implies that as students become readers, their scores on pre-reading tasks may not change or may even drop. Therefore, an evaluation of Albuquerque Reads or any reading intervention program should examine reading and pre-reading tasks separately.

> *H3: The impact of Albuquerque Reads does not become apparent until most or all of the intervention has taken place, at the end of the school year.*

In some cases, it is unwise to judge the effectiveness of a program before the entire intervention has taken place because improvement may not occur linearly over time. Rather, its effectiveness may not become apparent until the end of the intervention.

We cannot answer this research question definitively with the data currently available. However, at this time there seems to be less support for Hypotheses 1 and 2 than for Hypothesis 3.

Method

A thorough examination of these hypotheses requires DRA2 and, possibly, KDPR reliability studies as well as an examination of all KDPR Reading items through several test administrations, including DRA2 raw scores. One or more additional measures of reading ability would also be helpful checks on measurement validity. Unfortunately, these requirements go beyond the abilities of the data available for this study.

As a preliminary examination, however, we repeated the multivariate model described above for Table 5 to predict each individual KDPR Reading item for both winter and spring. The purpose

[9] Reading level as determined by the DRA2 is included in the 2008-09 KDPR Reading assessment. However, for the purposes of this research, the item is omitted when the two cohorts are combined to keep the two winter assessments equivalent. The DRA2 item from winter, 2009 is used in Table 6, however.

[10] While the hypothesis of biased test administration could apply to either DRA2 specifically or to all KDPR Reading items, in this case the concern has been raised mostly in regard to DRA2. In previous years, APS district personnel required teachers in Albuquerque Reads schools to administer DRA2 to all students in winter and spring, while teachers in other schools were advised to do so only if students were developmentally ready (i.e. only if reading). An informal re-testing of some Albuquerque Reads students in the 2006-07 SY revealed what some district experts believed to be a positive bias for these students' scores.

of this exercise is to gain a better understanding on how and when Albuquerque Reads impacts reading scores.[11]

Results

Table 6 lists the twelve items assessed by the KDPR Reading assessment in the winter and/or spring windows, in approximate chronological order of expected mastery.[12] Results for the winter window combine 2007-08 and 2008-09 cohorts, unless otherwise noted. Results for the spring window rely on the 2007-08 cohort only. "None" means the probit analysis found no statistically significant impact of Albuquerque Reads on the item after controlling for previous achievement and socio-economic factors. "Strong" means Albuquerque Reads increased the ordinal score by at least half a point,[13] while "Medium" means Albuquerque Reads increased the score by less than half a point, but still statistically significantly.

Table 6: Impact of Albuquerque Reads on Individual KDPR Items in Winter and Spring Test Administrations

	Winter	**Spring**
Identifies letters of the alphabet	none	none
Phonemic awareness of initial consonant sounds	none	*n/a*
Uses book handling skills	none	none
Uses picture clues to predict words & story outcomes	none	none
Uses pictures & written symbols to represent ideas & events	none	none
Phoneme Segmentation Fluency	none	Strong
Retells familiar story with beginning, middle, & end	none	none
Demonstrates rhyming words	none	none
Writes first & last name accurately with correct letter formations	none	Strong
Can write all upper & lower case letters correctly when dictated	none	none
Nonsense Word Fluency	none	Medium
Reads at grade level, DRA2 level 3 (Winter: 2008-09 cohort only)	Medium	Strong

In the winter window, Albuquerque Reads shows no impact on any winter KDPR Reading item except the last, 'reads at grade level,' for which it has a statistically significant impact of medium effect size. In the spring, Albuquerque Reads shows no impact for very early pre-reading skills, such as letter identification, book handling skills, and phonemic awareness. However, Albuquerque Reads does positively impact several later pre-reading skills as well as the ability to read. Albuquerque Reads shows a strong impact on phoneme segmentation fluency, name-writing, and reading at grade level, raising scores by at least half a point. The program also impacts nonsense word fluency to a lesser degree.[14]

[11] Since all items are based on 0-3 ordinal measures, probit analyses are used rather than OLS regression. Control factors include: fall KDPR Reading score, gender, well-performing ethnicities, poverty status, special education status, English language learners, language of assessment, statistical peer group, and cohort. Analyses are clustered by school. Complete results available by request.

[12] Sorted by a district literacy specialist.

[13] A half-point increase represents a very large effect since these items are measured on only a four point scale. In no case does Albuquerque Reads increase scores for any item by more than one point.

[14] We repeated this analysis for each cohort of students separately, with very similar results.

Hypothesis 1 implies that findings would show Albuquerque Reads improves DRA2 scores but have no impact on other KDPR Reading items. Because Albuquerque Reads shows impacts for test items other than that assessed with DRA2, these results imply that either Albuquerque Reads teachers are not more biased in their administration of DRA2 than other teachers (i.e. Hypothesis 1 is incorrect), or that teachers are biased in their administration of some other KDPR Reading items as well. So, while these results are not consistent the hypothesis of biased DRA2 test administration, they cannot rule out the possibility of biased test administration for all or additional KDPR items.

Nor do these results support Hypothesis 2, which accounts for the influence of Albuquerque Reads in the spring but not the winter by suggesting pre-reading skills do not improve and may diminish with increasing reading ability. Not only does Albuquerque Reads seem to improve some pre-reading skills, it does so even in the final KDPR window. However, it is puzzling that Albuquerque Reads appears to influence these pre-reading skills at the end of the year, but not beforehand.

Finally, these results are mostly consistent with the Hypothesis 3, which suggests the impact of Albuquerque Reads will be most apparent after most or all of the intervention has taken place. With the exception of reading ability in the winter, all other positive findings occur in the spring.

Summary

The results of this study and the evaluation of the 2006-07 cohort provide strong evidence that Albuquerque Reads students start the school year less academically prepared, yet score higher on DRA2 and other KDPR Reading items by the end of the kindergarten year. Albuquerque Reads students show more growth at the end of the year. By then, though overall KDPR Reading proficiency levels are equivalent, 53% of Albuquerque Reads students read at grade level, compared to 38% of their peers as gauged by the DRA2 reading level item.

Unfortunately, whether these difference are more apparent than real remains a question because of the possibility that Albuquerque Reads teachers are, knowingly or unknowingly, administering the DRA2 and possibly other assessment items with more positive bias than teachers in comparison schools. While the evidence is not strong, this analysis cannot rule out the hypothesis. More research is necessary to assess the hypothesis of teacher bias.

Limits to this Research

At least five important limitations exist to this research.

1. In light of the fact that the outcome measure used here may not be valid, perhaps the most important limitation is the lack of multiple outcome measures. While it is difficult to add assessments to teachers' already busy schedules, the addition of other outcome measures related to reading would go a long way towards assuring us of the validity of these results.
2. Albuquerque Reads differs from typical reading instruction in several ways, including more one-to-one instruction, more community involvement, more teacher-led small

group instruction, and more district support. This analysis cannot distinguish which of these differences or combination of differences might be responsible for the apparent success of Albuquerque Reads.

3. The generalizability of these findings is not strong. This analysis is limited to two cohorts, though results are similar to those of a third cohort published previously. All three cohorts consist of only three participating schools. Additional research in other settings with similar findings would strengthen our confidence in these results.
4. Other schools might employ "competing" reading strategies that focus on other grades or particular subgroups of students, such as struggling students. Therefore the outcome examined in this study, KDPR, may not be able to detect the success of other schools' programs.
5. Albuquerque Reads may be implemented with more fidelity than other schools' reading initiatives. While this analysis does not speak to the issue of fidelity, the fact that Albuquerque Reads is a well-established program, the ongoing training, and the attention paid to the program by district leaders and community members imply that fidelity is probably high. Many schools' reading initiatives (e.g. Wilson, Avenues, and K-3 Plus) have been implemented somewhat more recently and often with less district support.

Conclusion

This study points to three areas for future research. First, the district should continue to examine Albuquerque Reads outcome measures in a similar way in order to increase our confidence in the validity and generalizability of these results. At this point in time, the evidence showing the positive impact of Albuquerque Reads is limited to just two cohorts and three schools.

Second, it seems most important to examine the reliability (and therefore the validity) of DRA2 and possibly other KDPR Reading items as administered by teachers in Albuquerque Reads and peer schools. Although the Albuquerque Reads students seem to make better gains and proportionately more seem to be reading at grade level, we cannot rule out possible teacher bias, though certainly all teachers, not only Albuquerque Reads teachers, have an incentive to show their students are doing well. A reliability test would not only put to rest the question of Albuquerque Reads' impact, but would also inform the district on how much confidence we can have in reading results for our youngest students more generally.

Third, program and district personnel may consider incorporating formative evaluation techniques in the future. To date, research regarding Albuquerque Reads has not examined differences in results, fidelity, or effort among participating schools, for instance – information that could be very helpful particularly if the program is to expand. Implementation challenges, such as recruiting volunteers, could be examined in more detail. Finally, this study suggests that Albuquerque Reads may be particularly helpful in developing later pre-reading skills and reading ability. Further research may help determine, then, whether the program could best be used for students who have already mastered early pre-reading skills rather than all students.

By including many statistical control factors as well as a comparison group, this study has attempted to be a 'hard test' of Albuquerque Reads' effectiveness. The program provides strong evidence that the program improves reading ability. However, the program demands high

resources in terms of district and school personnel time as well as support materials. Therefore, ascertaining the reliability and validity of the measurements we use seems a most important next step.