Direct Instruction in the Guam Public Schools: An Analysis of Changes in Stanford Achievement Test Scores

Technical Report 2013-2

Jean Stockard, Director of Research
NATIONAL INSTITUTE FOR DIRECT INSTRUCTION
APRIL 29, 2013

Direct Instruction in the Guam Public Schools:
An Analysis of Changes in Stanford Achievement Test Scores
Executive Summary

This report examines changes in the average Stanford Achievement Test (SAT) scores in reading of first and fifth grade students in elementary schools in the Guam Public School System (GPSS) from Spring, 1993, through Spring, 2011. A systematic implementation of Direct Instruction, involving training and support, began in 24 schools in the fall of 2003, starting with the lower grades and gradually expanding to the upper grades. In the fall of 2008 15 of these schools ceased using DI, while the remaining schools and two additional newly created schools, continued to use the program.

The analysis focuses on differences in average SAT scores in schools and years when DI was implemented and when it was not implemented. A multiple base-line interrupted time series analysis, which is especially well suited to the analysis of aggregated data from multiple groups with varying characteristics and over a relatively long time span, supplements simple t-test comparisons. Three sets of results are presented: 1) an analysis of variation in first grade achievement using all years of data, 2) an analysis of variation in fifth grade achievement using all years of data, and 3) an analysis of fifth grade achievement limited to cohorts of students for whom data on first grade achievement were also available. To provide further checks on the results the analyses of first and fifth grade achievement were also conducted with a sub-sample that omitted extreme outliers. Differential rates of poverty in the schools were controlled in two ways: 1) by computing descriptive results for both the total sample and when the one low-poverty school was omitted and 2) by introducing the rate of free and reduced lunch as a control variable in the time series analysis.

The results are similar across all analyses. There were significantly higher achievement scores in school years in which students were exposed to Direct Instruction. Effect sizes are all above the level of .25, often used to denote educational importance, and range from .27 to .86 (mean = .43). These statistically significant and educationally important results appear both when controls are introduced for the poverty status of the schools and when they are not included. They appear with analyses of both first and fifth grade achievement and when the sample is limited to cohorts that had data for both grade levels. The results replicate other studies that have found that students exposed to Direct Instruction have higher achievement scores and that this impact occurs throughout the elementary years.

Direct Instruction in the Guam Public Schools: An Analysis of Changes in Stanford Achievement Test Scores

This report examines changes in the average reading Stanford Achievement Test (SAT) Scores of first and fifth grade students in elementary schools in the Guam Public School System (GPSS) from Spring, 1993, through Spring, 2011. The analysis focuses on the extent to which scores increased in the years and schools in which Direct Instruction programs were implemented. Sections below describe the schools in the study, the data and analysis techniques that were used, the results, and a short summary and discussion.

Schools in the Study

Data on school characteristics were provided through the GPSS, and Table 1 summarizes these characteristics at the end of the study period (2010-11). Information regarding each school is given in the top rows, and the minimum, maximum, and average values for each characteristic are in the bottom three rows. It can be seen that the schools varied widely in size, from only 54 students to almost 900, although the average was just over 500 students. Some schools had very high absentee rates, with a maximum of 15 percent, although the average was substantially less (6.7 percent). Variation was even greater in the extent to which schools suspended students, with rates varying from 22 percent in one school to none in another. The average rate of poverty was quite high. On average, across all the schools, 75 percent of students qualified for free or reduced lunch. Only one school (School 2) had a rate substantially less than the others, with only 46 percent of the students qualifying. One school (School 14) only enrolled students in grades K and 1. After grade 1, the students in School 14 attended School 25, which had no students in the two lowest grades. Thus, data on average achievement were available for 26 schools for both of the grades analyzed (all but School 25 for the analysis of grade 1 and all but School 14 for the analysis of grade 5).

In the fall of 2003, all but one of the 25 Guam elementary schools that were then in existence began systematically implementing Direct Instruction.[1] The exception was School 2, the school with the lowest level of poverty. As summarized in Table 2, the implementation began with the use of DI programs for reading with grades K-2. In 2004-05 this expanded to include reading instruction for students in grades K-5. In the fall of 2008 fifteen of the

[1] Some Guam schools had used Direct Instruction programs in previous years, but the implementations were neither systemic nor supported by technical assistance. Note that, to the extent that this occasional use raised test scores, it would have a conservative effect on the results, minimizing the reported impact of the systematic implementation of DI.

Table 1

Characteristics of Elementary Schools, Guam Department of Education, 2011

School	Enrollment Grades K-5	Average Absentee %	Free or Reduced Lunch %	Suspension Rate	DI Used in 2011
1	532	8.0%	78.9%	2.9%	yes
2	530	15.4%	46.0%	3.4%	no
3	485	6.2%	85.6%	8.3%	no
4	480	7.1%	63.1%	2.1%	yes
5	315	8.0%	94.3%	5.4%	no
6	567	7.7%	75.1%	3.3%	no
7	739	8.9%	74.6%	3.1%	yes
8	894	2.8%	81.4%	3.5%	no
9	54	3.0%	100.0%	14.8%	no
10	385	1.3%	59.0%	5.2%	no
11	250	10.4%	80.4%	0.7%	yes
12	755	3.8%	71.3%	2.3%	yes
13	523	9.4%	84.5%	21.6%	yes
14	289	8.4%	64.4%	0.0%	yes
15	633	6.8%	67.8%	2.1%	yes
16	680	0.5%	75.0%	0.8%	no
17	637	6.6%	72.4%	3.7%	yes
18	443	8.5%	86.7%	0.0%	no
19	437	7.8%	84.9%	1.3%	no
20	196	4.8%	81.1%	4.6%	no
21	492	7.4%	60.2%	0.4%	no
22	382	2.2%	84.6%	1.8%	no
23	634	8.1%	77.6%	7.0%	no
24	237	7.1%	79.3%	0.0%	no
25	598	5.1%	60.5%	1.7%	no
26	845	7.1%	76.0%	0.0%	yes
27	710	8.0%	74.9%	3.3%	yes
Minimum	54	0.5%	46.0%	0.0%	
Maximum	894	15.4%	100.0%	21.6%	
Average	508	6.7%	75.5%	3.8%	

Note: School 14 only had grades K-1, and School 25 only had Grades 2 and above. Students from School 14 transitioned to School 25 in second grade.

schools stopped using the program. Nine of the 24 schools continued to use DI through 2011, and two newly created schools (one opening in 2008-09 and one opening in 2009-10) also adopted the program. The last column of Table 1 indicates whether the school used Direct Instruction in 2010-11. In 2011 these DI schools were slightly larger than the other schools and had slightly greater absentee rates, but only small differences, overall, in the receipt of free or reduced lunch or use of suspension.

Table 2
Summary of Implementation Patterns

		DI Programs Implemented		
# of Schools	Year	DI Read K-2	DI Read K-5	no DI
25	1993-03	0	0	25
25	2003-04	24	0	1
25	2004-05	24	24	1
25	2005-06	24	24	1
25	2006-07	24	24	1
26	2008-09	8	8	16
27	2009-10	9	9	16
27	2010-11	10	10	16

Note: 2007-08 is omitted because no data were provided for that year.

Training and implementation support were provided by the National Institute for Direct Instruction (NIFDI). A five day in-service for all teachers, administrators, and teaching assistants was held in the summer of 2003, before the first year of implementation. Beginning the same year a Guam-based Direct Instruction specialist provided general support and "trouble shooting" services to all the schools, working with building administrators to help support the implementations at their sites. In addition, schools were provided consulting services to help teachers implement the program with greater fidelity within their classrooms. However, the amount of support varied across the district. Six schools received twenty days each year in direct consultation services, while the others received substantially less (six days annually).

Table 3 provides a summary of schools' years of DI implementation and level of support, grouping the 27 schools into six categories: 1) One school (Number 2) never used Direct Instruction; 2) Thirteen schools used DI from 2003-04 through 2007-08 and received six days each year of external support; 3) Two schools used DI from 2003-04 through 2007-08 and received twenty days each year in external support; 4) Five schools have used DI since

2003-04 and received 6 days per year of external consultation through the study period; 5) Two schools opened recently (one in the fall of 2008 and one in the fall of 2009) and have always used DI; and 6) Four schools have used DI since 2003-04 and received 20 days per year of external support. Schools in groups 4-6 were still using DI in 2010-11, while those in groups 2-3 were not. The school in group 1 never used DI. In total, DI was implemented in 35 percent of the years in the first grade analysis and 22 percent of the years in the fifth grade analysis. Students had been exposed to Direct Instruction since kindergarten in 29 percent of the school years in the first grade analysis, but only three percent of the school years in the fifth grade analysis.

Table 3
Summary of Schools' Experiences with Direct Instruction

1. Never Used Direct Instruction: One School
2. Used Direct Instruction 2003-04 through 2007-08, 6 days/year of consultation: Thirteen Schools
3. Used Direct Instruction 2003-04 through 2007-08, 20 days/year of consultation: Two Schools
4. Used Direct Instruction since 2003-04, 6 days/year of consultation : Five Schools
5. Always used Direct Instruction, 6 days/year of consultation: Two Schools, one opening in 2008-09 and one opening in 2009-10
6. Used Direct Instruction since 2003-04, 20 days/year of consultation: Four Schools

The NIFDI personnel involved with training and technical implementation support reported varying levels of commitment and implementation of the Direct Instruction model across the schools in the system. For instance, some administrators did not attend the pre-service sessions, and the extent of administrative support and acceptance of the approach was reported to vary from one site to another. An independent consulting firm examined the implementation in 2008 (Robinson, Towner, Caros, & Billups, 2008). Based on multiple observations they reported ratings of schools in three areas: 1) the quality of implementation, including the quality of school leadership, following the model, accountability, stability of administration and reading coordinators, and quality of peer coaching; 2) the quality of instruction, reflecting acceleration of students, accountability, and teaching to mastery; and 3) the quality of lessons, defined as appropriate pacing and lesson coverage given the time allotted. Results were reported for the end of the third, fourth, and fifth years of implementation (2005-06 to 2007-08) for the quality of implementation and instruction and for the fourth and fifth years for the quality of lessons. Schools were assigned a numeric score indicating whether they were in the top third, middle third, or bottom third relative to other schools in the district in a given area and year. The evaluators stressed that a top ranking should not be construed as indicating that a school met all the model requirements, but simply that their performance was better than that of other schools in the system. Fidelity data were not available for the three schools, including the two that

most recently began their DI implementations (numbers 1 and 15 in Table One) and one other (number 8 in Table One). The scores were highly correlated across the three dimensions and over time. Thus, for this analysis the fidelity scores across areas and years were averaged and schools with average scores greater than 2.5 (n=6) were compared with those with lower scores (n=17). Schools that received more extensive technical support tended to have higher fidelity ratings and were more likely to retain Direct Instruction. An extended discussion of these relationships is included in a companion report (Stockard, 2013).

Data and Analysis Procedures

Summary data on average Spring SAT scores for schools in the system were used in the analysis. For 1993 to 2007 data were provided in the form of percentiles in spreadsheet format. For 2009 to 2011 data were obtained from graphs included in the Superintendent's annual reports, where they were given as scale scores. The scale scores were transformed to percentiles using the SAT Spring Multilevel Norms Book (Pearson, 2003). The standard formula for transforming percentiles to Normal Curve Equivalent (NCE) scores was used, with the calculations completed in Excel and Stata:

Invnorm (percentile) * 21.06 + 50 (1)

By definition, NCE scores are normally distributed with a mean of 50 and a standard deviation of 21.06. Unlike percentiles, NCE scores are appropriate for statistical manipulations.

While the data represent results from both the Stanford 10 and the Stanford 9, scores from the two versions were constructed to ensure equivalence (Pearson, 2003, p. 22). Information on the percentage of students receiving free or reduced lunch within each school was only supplied for 2011, and these data were used as a proxy for socio-economic status (See Table 1).

As noted above, no data were available for Spring, 2008. In addition, data were not available for each school and year because some schools were not in existence for the entire time-frame and/or no data were given for some schools for some years. Data were available for all 18 years for 10 schools, 17 years for 1 school, 15 years for another, 13 years for 3 schools (starting in 1998), 11 years for one (starting in 2000), 5 years for 8 schools (starting in 2007), 4 years for one school (beginning in 2007, but missing data for 2010), 3 years for 1 (beginning in 2009), and 2 years for another (beginning in 2010). In addition, as explained in the discussion of results below, the availability of data varied somewhat from one grade level to the other. The analysis technique used is robust against such patterns of missing data.

Based on previous research it was expected that schools would have higher achievement scores when they implemented Direct Instruction. It was also expected that this impact would be greater 1) after the new program had been fully implemented and students had been exposed to it since the beginning of their schooling experience; and 2) when the program was implemented with greater fidelity (Benner, Nelson, Stage, & Ralston, 2010; Carlson & Francis, 2002; Gersten, Carnine, & Williams, 1982; Gersten, Carnine, Zoref, & Cronin, 1986; Ross et al., 2004; Stockard, 2011). Decades of research have also shown that schools with greater concentrations of poverty have lower achievement scores, and it was important to control for this influence in any analysis.

The analysis began by looking at the data for first and fifth grade separately. Data were first examined descriptively and with simple inferential statistics, using the average score for each school and year as the unit of analysis. Means and standard deviations of average achievement scores in school years with and without DI were compared using t-tests and effect sizes (Cohen's d). As a way to control for the impact of poverty, the analysis was conducted with both the total sample and then omitting the school whose poverty rate was substantially lower than the other schools. (This school never implemented Direct Instruction.) Because the analysis of descriptive data indicated a wide range of scores, the calculations were also conducted with a sub-sample of school-years that included only the middle 80 percent of scores (omitting cases below the 10th and about the 90th percentile), thus removing the potential impact of outliers.

Data were subjected to a multivariate analysis using a multiple base-line interrupted time-series approach. Several authors have suggested that this technique may be especially suited to long-term evaluations of community or large-group interventions (Biglan, Ary, & Wagenaar, 2000; Biglan, Flay, Komro, Wagenaar, & Kjelistrand, 2012; Madigan & Cross, 2012). In the present analysis a mixed model approach was used, in which it was assumed that yearly assessments (i) were nested within each school (j). Maximum likelihood estimates were used, with random intercepts and controls for clustering of observations within school, employing the xtmixed procedure in Stata. Six increasingly more complex models were tested, paralleling the expectations noted above:

An intercept-only, or baseline model, which only included the random effects of school. This is equivalent to an analysis of variance model with schools as the factor and can be written as

$$Y_{ij} = \beta + \zeta_j + \varepsilon_{ij} \qquad (2),$$

where Y_{ij} is a score for year i in school j, β (beta) is the overall mean, ζ_j (zeta) is the error term (or residual variance) associated with each school j (the difference of the school mean from the overall mean), and ε_{ij} (epsilon) is the error term associated with each observation (the difference of each observation from the mean).

A model that added the percentage of students receiving free or reduced lunch as a predictor,

$Y_{ij} = \beta + \beta_1(X_{1j}) + \zeta_j + \varepsilon_{ij}$, (3),

where X_{1j} is the percentage of students receiving free or reduced lunch in each school j and β_1 is the regression coefficient, or fixed effect coefficient, associated with this variable.

A model that added a dummy variable indicating whether or not the school had implemented Direct Instruction, with a value of 1 if the school j had DI in a given year and 0 if it had some other program:

$Y_{ij} = \beta + \beta_1(X_{1j}) + \beta_2(X_{2ij}) + \zeta_j + \varepsilon_{ij}$, (4),

where X_{2ij} is the dummy variable indicating instructional program in year i and school j, and β_2 is the regression coefficient, or fixed effect coefficient, associated with this variable.

A model that added a dummy variable indicating the extent to which the model had been fully implemented and students would have, theoretically received the program since their kindergarten years. For instance, for the analysis of first grade results, the variable equaled 1 for the years in which the school had used DI for at least two years (since the students were in kindergarten) and zero for other years.

$Y_{ij} = \beta + \beta_1(X_{1j}) + \beta_2(X_{2ij}) + \beta_3(X_{3ij}) + \zeta_j + \varepsilon_{ij}$, (5),

where X_{3ij} is the dummy variable indicating full implementation and β_3 is the regression coefficient, or fixed effect coefficient, associated with this variable.

A model that added a dummy variable indicating higher fidelity of DI implementation, distinguishing schools with average scale scores of 2.5 or greater from other schools, as described above, to Model 3.[2]

$Y_{ij} = \beta + \beta_1(X_{1j}) + \beta_2(X_{2ij}) + \beta_4(X_{4ij}) + \zeta_j + \varepsilon_{ij}$, (6),

where X_{4ij} is the dummy variable indicating that a school's implementation of DI was of relatively higher quality than other schools in the system and β_4 is the regression coefficient, or fixed effect coefficient, associated with this variable.

A model that included both the dummy variable indicating fuller implementation and the dummy variable indicating higher fidelity.

$Y_{ij} = \beta + \beta_1(X_{1j}) + \beta_2(X_{2ij}) + \beta_3(X_{3ij}) + \beta_4(X_{4ij}) + \zeta_j + \varepsilon_{ij}$, (7).

The -2 Log Likelihood statistics were used to determine the relative fit of the models and determine which models to examine further. Fixed effect coefficients and associated probability levels of the best models were examined to determine the relative impact of implementation related factors and socio-economic status on achievement. The coefficients

[2] The schools without fidelity measures were assigned to the zero category to differentiate them from the schools with known higher fidelity.

associated with the random effects were also used to calculate the percentage of variation between schools (called the correlation ratio) and the extent to which this variation was accounted for by the variables in the model.

Additional analyses replicated the results reported below. A standard time-series analysis, using xtregar in Stata, which allows the computation of an autocorrelation term, produced almost identical results. The impact of DI was usually slightly larger in these analyses, but the standard error was also larger. Because the mixed model approach also controls for autocorrelation (through the random effects and clustering options) and because the smaller standard errors indicated greater accuracy, those results are reported here.

To supplement the analyses with each year, a sub-group of cases restricted to cohorts for whom there were data for both first and fifth grade was analyzed. In other words, the analysis examined only the pairs of school years for which data were available for the same group of students in both first and fifth grade (e.g. first graders in 1999 who were fifth graders in 2003, first graders in 2000 who were fifth graders in 2004, etc.) Again, descriptive data and multivariate models were computed using the multiple base-line interrupted time-series approach. Six models, each more complex than the previous one and all having fifth grade average reading achievement as the dependent variable, were examined: 1) an intercept only base-line model, 2) a model that added the percentage of students receiving free or reduced lunch in the school, 3) a model that added the average first grade reading achievement of the cohort, 4) a dummy variable indicating whether or not the cohort had DI in their school during first grade, 5) a dummy variable indicating whether or not the cohort had DI in their school during fifth grade, and 6) two dummy variables indicating whether the DI implementation had higher fidelity in each of the grades.

It is important to note limitations to this analysis. While data were available for a relatively long span of time, they were not available for all years in the data period or for all schools. In addition, no independent knowledge of the fidelity of the testing procedure or scoring was available. Similarly, the measure of implementation fidelity was obtained from a secondary source and involved relatively crude groupings of schools and no measures by grade level. The measure of school poverty level was obtained for only one year. The most important limitation was, of course, the availability of only aggregate data for each school rather than student- level data. Such average scores reflect the achievement of students with varying levels of exposure to the model, including those who had been in a school throughout their career and also those who potentially transferred from other schools with alternative models. In addition, it precludes the use of individual level control variables. Nevertheless, the analysis is based on a data set not unlike that available to many schools and communities and has involved numerous checks for accuracy and consistency.

Direct Instruction and First Grade Achievement

Data on first grade average SAT scores were available for 306 school years, covering the period 1993 through 2011. Table 4 reports NCE scores for school years when students were exposed to DI and when they were not. DI was implemented for 35 percent of the school years for which data were available. Students had experienced DI since kindergarten in 29 percent of the school years, and high fidelity was reported for the schools in 13 percent of the years. Two sets of data are given for school years with no exposure to DI: one that includes all schools and another that omits the school with the relatively low poverty rate. Also included are results of t-tests and the associated effect sizes. The left hand columns give results for the total group, and the right hand columns give results for the reduced sample, removing data points below the 10th and above the 90th percentile.

Table 4

Average Reading SAT Scores (NCE) by Presence of Direct Instruction, Full and Reduced Sample, First Grade

	Full Sample			Reduced Sample		
Descriptive Statistics	Mean	S.D.	N	Mean	S.D.	N
Direct Instruction	46.41	4.82	106	45.75	3.74	94
No DI (Total group)	44.60	6.93	200	44.63	4.01	149
No DI (Sch. 2 omitted)	43.95	6.58	182	44.25	3.95	135
Comparisons	t-ratio	prob.	ef. size	t-ratio	prob.	ef. size
DI vs. No DI (Total group)	2.41	0.008	0.29	2.19	0.01	0.29
DI vs. No DI (Sch. 2 omitted)	3.37	<.001	0.40	2.9	0.002	0.38

Note: N refers to the number of school years for which data were available. One-tail probability levels were used to match the expectation that Direct Instruction school-years would have higher levels of achievement. The reduced sample omits outliers, defined as school-years of data that fell in the top 90th and bottom 10th percentiles of the total distribution.

It can be seen that average achievement scores were higher for schools and years in which DI was used as the curriculum than when other approaches were used. Standard deviations were also smaller in the years that DI was used, indicating more homogeneity (fewer students scoring at extraordinarily low and/or high levels). All results are statistically significant, and, in all comparisons, effect sizes (Cohen's d) are equal to or higher than the level of .25 typically used to denote educational importance (Tallmadge, 1977). As would be expected, the magnitude of the differences is stronger when the low poverty school is omitted. Effect sizes are also stronger with results with the reduced sample, primarily because the standard deviation is smaller when outliers are removed. On average, however, none of the values surpasses the national mean of 50. The average NCE score for the total group of DI schools, as reported in Table 4, is equivalent to the 43rd percentile; for the non-DI school years the average NCE scores are equivalent to the 39th and 40th percentiles.

Tables 5 and 6 give results for the mixed model time series analysis, including controls for the clustering of schools, poverty status of the schools, fidelity of implementation, and exposure of students to the DI model. Table 5 summarizes the model fit statistics for the analyses. The -2 Log Likelihood ratios have a chi-square distribution and can be compared from one model to another to determine if the variables added in a model provide better fit to the data. Lower values indicate a better fit. Results are reported for both the total and reduced (without outliers) samples. In both analyses, models 2 and 3 are better fitting than the less complex models. Models 4, 5 and 6, which add the measures of having DI since kindergarten and of high fidelity, do not provide a significantly better fit. In other words, for

Table 5

Model Fit Statistics First Grade Achievement Analysis of Reading, Full and Reduced Samples

Reading - Full Sample					
	LL Ratio	-2 LL	-2 LL change	df	prob.
Model 1	-991.15	1982.29			
Model 2	-987.42	1974.84	7.46	1	<.01
Model 3	-982.24	1964.49	10.35	1	<.01
Model 4	-982.24	1964.48	0.01	1	n.s.
Model 5	-981.92	1963.84	0.65	1	n.s.
Model 6	-981.92	1963.84	0.00	2	n.s.
Reading - Reduced Sample (Outliers Omitted)					
	LL Ratio	-2 LL	-2 LL change	df	prob.
Model 1	-672.24	1344.48			
Model 2	-667.63	1335.26	9.22	1	<.01
Model 3	-662.23	1324.46	10.80	1	<.01
Model 4	-661.89	1323.78	0.68	1	n.s.
Model 5	-662.2	1324.40	0.06	1	n.s.
Model 6	-661.87	1323.74	0.72	2	n.s.
Variables in the Models					
			Models		
	2	3	4	5	6
FRL	X	X	X	X	X
DI		X	X	X	X
DI since K			X		X
High Fid				X	X

Note: The degrees of freedom represent the number of new variables added to a model. Change in the -2 LL values for models 4, 5, and 6 were calculated by comparing to Model 3.

both analyses, Model 3, which includes the indicator of poverty status and use of DI provides the best fit to the data.

Table 6 gives the fixed effects coefficients associated with Model 3 for each of the analyses. As expected, schools with higher percentages of students receiving free and reduced lunch have significantly lower achievement scores (see the negative coefficient associated with FRL). In addition, as expected, the coefficient associated with Direct Instruction is positive and significant in all analyses. In the years in which schools used DI the average first grade achievement score is significantly greater than in other years.

Table 6

Mixed Model Fixed Effects and Random Effects Coefficients, First Grade Reading and Mathematics Achievement, Full and Reduced Sample

Full Sample			
		Reading	
Fixed Effects	coef.	s.e.	prob.
Constant	53.6	3.50	<.001
Free or Reduced Lunch %	-0.13	0.05	0.01
Direct Instruction	2.45	0.74	0.001
Random Effects			
Between Schools	1.94	0.52	<.001
Within Schools (Residual)	5.80	0.24	<.001
Reduced Sample			
		Reading	
Fixed Effects	coef.	s.e.	prob.
Constant	51.32	2.43	<.001
Free or Reduced Lunch %	-0.10	0.03	0.003
Direct Instruction	1.75	0.50	<.001
Random Effects			
Between Schools	1.38	0.40	<.001
Within Schools (Residual)	3.53	0.17	<.001

Note: For the analysis with the full sample 26 schools were included with an average of 11.8 observations per school, a minimum of 2 observations and a maximum of 18 and a total of 306 observations. Results from the intercept only model (Model 1), indicated a between schools s.d. of 2.15 and a within schools s.d. of 5.95 for reading. The corresponding correlation ratio (proportion of variation between schools) was .12. For the analysis with the reduced sample there were 26 schools included, with a total of 243 observations. The number of observations per school ranged from 2 to 16 for reading, with an average of 9.3. Results from the intercept only model (Model 1), indicated a between schools s.d. of 1.46 and a within schools s.d. of 3.68. The correlation ratio was .14.

The random effects coefficients can be used to calculate the correlation ratio and an indicator of the proportionate reduction of error from the baseline to Model 3. The correlation ratio, or percentage of the total variance that is between schools, equals .12 for the full sample and .14 for the reduced sample that omitted the outliers. The proportionate reduction of error measure describes the percentage of variation between schools that is accounted for by the variables in the model and equals .19 for the full sample and .11 for the reduced sample.[3]

The fixed effect coefficients associated with Direct Instruction can be used to calculate effect sizes, showing the effect of implementing DI net of the school poverty rate and within-school variation. The calculations show that controlling for the poverty level of the schools results in a substantially higher effect size for reading (d = .39 for the full sample and d = .44 for the reduced sample), an increase by about 1/3 over the value when SES is not controlled (d = .29, the value shown in Table 4).[4]

Fifth Grade Achievement

There was somewhat less data for fifth grade achievement than for first grade, with no data reported for 2006 or 2010. (Recall that there were no data for 2008 for either grade.) There were a total of 259 school years of data for analysis for fifth graders, about 85% of the number of years available for first grade. In addition, because DI was implemented at later years a smaller proportion of the school years (22 percent) involved DI implementations. Only three percent of the school years for the fifth grade analysis involved cohorts that had experienced DI since kindergarten, and only eight percent involved implementations with high fidelity.

Table 7 gives descriptive statistics (means and standard deviations) of NCE scores for fifth graders for school years when students were exposed to DI and when they were not. As with the analyses for first grade, two sets of data are given for school years with no exposure: one that includes all schools and another that omits the school with the relatively low poverty rate. Results are also given for the full sample and the reduced sample, which omits the outliers (school years with scores above the 90th percentile and below the 10th percentile); and t-test values and associated effect sizes are included. Again, scores are significantly higher in schools and years when DI was implemented, and the results are similar for both the full sample and the reduced sample. As with the analysis of first grade, the effect size increases when the low poverty school is omitted.

---

[3] The correlation ratio is calculated by dividing the between schools variance (calculated by squaring the standard deviation given in Table 6) by the total variance (the sum of the between schools and within schools variance). The proportionate reduction in error is calculated by dividing the difference of the between schools variance in model 1 and model 3 by the between schools variance in model 1.

[4] The adjusted effect size is calculated by dividing the coefficient associated with DI by the common standard deviation.

Table 7

Average Reading Achievement Scores (NCE) by Presence of Direct Instruction, Full and Reduced Sample, Fifth Grade

	Full Sample			Reduced Sample		
	Mean	S.D.	N	Mean	S.D.	N
Direct Instruction	36.23	3.33	57	35.67	2.62	49
No DI (Total group)	34.17	4.54	202	34.65	2.68	152
No DI (Sch. 2 omitted)	33.71	4.19	186	34.61	2.69	146
Comparisons	t-ratio	prob.	ef. size	t-ratio	prob.	ef. size
DI vs. No DI (total group)	3.19	0.001	0.47	2.32	0.01	0.38
DI vs. No DI (Sch. 2 omit.)	4.16	<.001	0.61	2.40	0.01	0.39

Note: N refers to the number of school years for which data were available. One-tail probability levels were used to match the expectation that Direct Instruction school-years would have higher levels of achievement.

Table 8 gives the model fit statistics. In both analyses, the best fitting model is Model 3, including the measure of school poverty and DI implementation, paralleling the results with the first grade data. Table 9 gives the coefficients associated with these best fitting models. The results parallel those found with the t-tests, with schools having significantly higher fifth grade reading achievement scores in years when DI was implemented. The magnitude of the coefficients is slightly smaller in results with the reduced sample than in the full sample, but the coefficients are still statistically significant. The effect sizes associated with having DI, as calculated from the models, are .86 for the full sample and .43 for the sample that omitted outliers, both well above the level usually used to denote educational importance.

The correlation ratios and the PRE measures are substantially higher for the fifth grade results than for the first grade results, indicating that differences between schools are more marked at the higher grade level. The correlation ratio (percentage of total variation that was between schools) is .28 for the full sample and .37 for the reduced sample. The proportionate reduction in error from the baseline model to Model 3 is .72 for the full model and .59 for the reduced model. In other words, over a fourth of the variation in fifth grade reading achievement is between schools and over half of these between-school differences can be explained by the two variables in Model 3: the poverty level of the school and receipt of Direct Instruction.

Table 8

Model Fit Statistics Fifth Grade Achievement Analysis

Full Sample					
	LL Ratio	-2 LL	-2 LL change	df	prob.
Model 1	-725.8	1451.5			
Model 2	-714.5	1429.0	22.52	1	<.001
Model 3	-706.3	1412.6	16.40	1	<.001
Model 4	-706.3	1412.5	0.05	1	n.s.
Model 5	-706.3	1412.6	0.0004	1	n.s.
Model 6	-706.3	1412.5	0.05	2	n.s.
Reduced Sample (Without Outliers)					
	LL Ratio	-2 LL	-2 LL change	df	prob.
Model 1	-463.17	926.34			
Model 2	-454.57	909.14	17.2	1	<.001
Model 3	-449.86	899.72	9.42	1	<.01
Model 4	-449.86	899.72	0.00	1	n.s.
Model 5	-449.74	899.48	0.24	1	n.s.
Model 6	-449.74	899.48	0.24	2	n.s.
Variables in the Models					
	Models				
	2	3	4	5	6
FRL	X	X	X	X	X
DI		X	X	X	X
DI since K			X		X
High Fid				X	X

Note: The degrees of freedom represent the number of new variables added to a model. Change in the -2 LL values for models 4, 5 and 6 were calculated by comparing to Model 3.

Cohort Analysis

As a further check on the results the sample was restricted to a subgroup that included only pairings of years that would theoretically include the same groups of students in first and fifth grade. Table 10 lists the cohorts that are used in the analysis, beginning with the cohort in first grade in spring 1995 and fifth grade in spring 1999, through the cohort in first grade in 2007 and in fifth grade in 2011. Numbers in the table report the number of schools in each category of implementation status and cohort. For instance, there are 12 school

Table 9

Mixed Model Fixed Effects and Random Effects Coefficients, Fifth Grade Reading Achievement, Full and Reduced Sample

Full Sample			
Fixed Effects	coef.	s.e.	prob.
Constant	47.15	2.24	<.001
Free or Reduced Lunch %	-0.17	0.03	<.001
Direct Instruction	2.32	0.56	<.001
Random Effects			
Between Schools	1.21	0.31	<.001
Within Schools (Resid.)	3.57	0.16	<.001
Reduced Sample (Without Outliers)			
Fixed Effects	coef.	s.e.	prob.
Constant	43.55	1.87	<.001
Free or Reduced Lunch %	-0.12	0.02	<.001
Direct Instruction	1.18	0.38	0.002
Random Effects			
Between Schools	1.07	0.26	<.001
Within Schools (Resid.)	2.13	0.11	<.001

Note: For the analyses of the total group there were 259 observations. Twenty-six schools were included with an average of 10 observations per school, a minimum of 1 observations and a maximum of 16. Results from the intercept only model (Model 1), indicated a between schools s.d. of 2.30 and a within schools s.d. of 3.71. The corresponding correlation ratio (proportion of variation between schools) was .28. For the analyses of the reduced sample there were 201 observations. The observations came from 26 schools with from 1 to 15 observations per school and an average of 7.7 observations per school. The between school standard deviation for the base-line, intercept only, model was 1.67 for reading; the within school s.d. was 2.20, and the corresponding correlation ratio was .37.

cohorts that were in first grade in the spring of 1995 and fifth grade in the spring of 1999 and none of these cohorts had DI in either first or fifth grade. There are 16 school cohorts that were in first grade in 2005 and fifth grade in 2009. One of these (School 2) never had DI, 12 had DI in first grade but not in fifth, and 3 had DI in both grades. Note that no data are available for 3 possible cohorts (first grade in 2002, 2004 and 2006 corresponding to fifth grade in 2006, 2008 and 2010), resulting in a total of 153 school years in the analysis. It must also be stressed that, because the only data available are average scores of students in the school, there is no way of judging the extent to which the average values are

obtained from the same group of students. That is, there is no correction or control for differential mobility between the schools.[5]

Table 10
Cohorts in the Analysis

Years Gr 1, Gr 5	No DI	First Only	Fifth Only	Both 1 and 5	Total
1995, 1999	12	0	0	0	12
1996, 2000	12	0	0	0	12
1997, 2001	12	0	0	0	12
1998, 2002	15	0	0	0	15
1999, 2003	15	0	0	0	15
2000, 2004	16	0	0	0	16
2001, 2005	1	0	15	0	16
2003, 2007	1	0	15	0	16
2005, 2009	1	12	0	3	16
2007, 2011	1	14	0	8	23
Total (n)	86	26	30	11	153
Total (%)	56%	17%	20%	7%	100%

Note: Numbers in the table report the number of schools in each category of implementation status and cohort. For instance, there were 12 school cohorts that were in first grade in the spring of 1995 and fifth grade in the spring of 1999 and none of these cohorts had DI in either first or fifth grade. There were 16 school cohorts that were in first grade in 2005 and fifth grade in 2009. One of these (School Number 2) never had DI, 12 had DI in first grade but not in fifth, and 3 had DI in both grades.

Table 11 reports the descriptive data on average reading scores in fifth grade for the four groups listed in Table 10. The highest averages appeared for students who had DI in both first and fifth grade, followed by those with DI in grade 5 only and the total group including the low poverty school. As expected, the average score dropped substantially when the low poverty school was omitted from the total group. The lowest score occurred for the schools that had DI only in first grade, having dropped the program by the students' fifth grade year. The analyses of variance comparing across all four types of schools were not statistically significant, as summarized in the two bottom rows of the first panel. The analyses of variance were supplemented with t-ratios comparing each pair of schools; and these results are given in the bottom panel of Table 11 along with the associated effect sizes (Cohen's d). The only statistically significant result involves the comparison of schools with DI only in first grade and those that had the program in both first and fifth grade. This comparison also had the largest effect size (.62). In other words, these results indicate that the highest average fifth grade achievement scores occur in schools and cohorts that had DI in both their first

[5] In general, student mobility tends to be positively associated with school poverty; that is, schools with higher rates of free or reduced lunch generally have higher rates of mobility. The inclusion of free and reduced lunch in the time series models thus helps control for this influence.

and fifth grade years; the lowest scores occur in schools and cohorts that employed DI when the students were in first grade but dropped the program by the time the students were in fifth grade.

Table 11

Average Fifth Grade Reading Achievement Scores (NCE) by Presence of Direct Instruction in First and Fifth Grade, Cohort Sample

	Mean	S.D.	N
Direct Instruction, Grade 1 only	34.69	3.07	26
Direct Instruction, Grade 5 only	35.98	3.59	30
Direct Instruction, Grade 1 and 5	36.59	2.64	11
No DI (Total group)	36.01	4.07	86
No DI (Sch. 2 omitted)	35.24	3.50	76
Comparisons	F-ratio	prob.	df
Total Group	1.04	0.38	3, 149
Omitting Sch. 2	1.19	0.39	3, 139
Pair-Wise Comparisons			
	t-ratio	prob.	eff size
DI Gr 1 only v. no DI (total group)	-1.53	0.94	-0.34
DI Gr 5 only v. no DI (total group)	-0.03	0.97	-0.01
DI Gr 1 and 5 v. no DI (total group)	0.46	0.32	0.15
DI Gr 1 only v. no DI (Sch. 2 omitted)	-0.72	0.76	-0.16
DI Gr 5 only v. no DI (Sch. 2 omitted)	0.98	0.17	0.21
DI Gr 1 and 5 v. no DI (Sch. 2 omitted)	1.23	0.11	0.39
DI Gr 1 only v. DI Gr 5 only	1.44	0.08	0.38
DI Gr 1 only v. DI Gr 1 and 5	1.79	0.04	0.62
DI Gr 5 only v. DI Gr 1 and 5	0.51	0.31	0.18

Note: N refers to the number of school years for which data were available. All probability levels are one-tail based on the hypotheses that scores would be higher with Direct Instruction for the comparisons with no DI, and for more recent exposure to the curriculum, for the comparisons between DI instructional categories.

Table 12 gives model fit statistics for each of the models. The analysis of changes in the fit statistics indicates that adding the measure of school poverty (Model 2) produces a significantly better fit, while adding achievement in first grade or having DI in first grade does not improve fit (Models 3 and 4). The addition of having DI in grade 5 (Model 5) significantly improves the model fit, but adding the measures of implementation fidelity (Model 6) does not. Thus, Model 5 was selected for further analysis.[6]

[6] The measures of first grade program and achievement were retained to provide additional controls for characteristics of the school. Results are virtually identical when these measures are omitted from the analysis.

Table 12

Model Fit Statistics Cohort Analysis of Reading Achievement

Full Sample					
	LL Ratio	-2 LL	-2 LL change	df	prob.
Model 1	-382.0	764.0			
Model 2	-369.3	738.6	25.34	1	<.001
Model 3	-369.3	738.6	0.00	1	n.s.
Model 4	-369.3	738.6	0.00	1	n.s.
Model 5	-367.3	734.6	4.00	1	<.05
Model 6	-366.6	733.1	1.52	2	n.s.
Variables in the Models					
		Models			
	2	3	4	5	6
FRL	X	X	X	X	X
Ach. Gr. 1		X	X	X	X
DI Gr. 1			X	X	X
DI Gr. 5				X	X
High Fid Gr 1, 5					X

Note: The degrees of freedom represent the number of new variables added to a model. Each model is compared the previous model.

Table 13 gives the fixed and random effect coefficients associated with Model 5. The fixed effect coefficients (in the top rows) show that average fifth grade achievement scores are significantly lower in schools that had higher levels of free or reduced lunch, but are significantly higher in schools that had Direct Instruction in fifth grade. Neither the average levels of first grade achievement nor having DI in first grade significantly influences the average scores in fifth grade. The effect size associated with having DI in fifth grade is .27, surpassing the usual criterion used to denote educational importance.[7] Analysis of the random effect coefficients (in the bottom rows and footnote to the table) indicates that over half of the total variation (52 percent) in average fifth grade achievement is between schools. The proportionate reduction of error is .79, indicating that this between-school variation is reduced by almost 80 percent by the variables within the model. In other words, four-fifths of the variation between schools in the average achievement of fifth graders can be explained by their rate of free and reduced lunch and whether or not the students had DI in fifth grade.

[7] As in the earlier analyses, the effect size is calculated by dividing the coefficient associated with having DI in fifth grade (1.023) by the standard deviation of the dependent variable (3.74).

Table 13

Mixed Model Fixed Effects and Random Effects Coefficients, Fifth Grade Reading Achievement, Cohort Analysis

Fixed Effects	coef.	s.e.	prob.
Constant	50.940	2.789	<.001
Free or Reduced Lunch %	-0.207	0.029	<.001
First Grade Average Achievement	0.003	0.031	0.936
DI in First Grade	-0.036	0.490	0.942
DI in Fifth Grade	1.023	0.471	0.030
Random Effects			
Between Schools	1.221	0.319	<.001
Within Schools (Resid.)	2.496	0.153	<.001

Note: There were 259 observations representing 23 schools, with an average of 6.7 observations per school, a minimum of 1 observations and a maximum of 10. Results from the intercept only model (Model 1), indicated a between schools s.d. of 2.66 and a within schools s.d. of 2.55 The corresponding correlation ratio (proportion of variation between schools) was .52. The between school variation was reduced by 79 percent from the base line to full model.

Summary and Discussion

This report examines changes in the average Stanford Achievement Test scores in reading of first and fifth grade students in elementary schools in the Guam Public School System from Spring, 1993, through Spring, 2011. A systematic implementation of Direct Instruction, involving training and support began in 24 schools in the fall of 2003, starting with the lower grades and gradually expanding to include instruction in the upper grades. In the fall of 2008 15 of these schools ceased using DI, while the remaining schools and two additional newly created schools, continued to use the program.

The analysis examines variations in the average SAT reading scores of schools over this time period and the extent to which having Direct Instruction can account for these variations. Three sets of results are presented: 1) an analysis of variation in first grade achievement using all school years, 2) an analysis of variation in fifth grade achievement using all school years, and 3) an analysis of fifth grade achievement limited to cohorts of students for whom data on first grade achievement are also available. To provide further checks on the results the analyses of first and fifth grade achievement are conducted with all schools and years and with a sub-sample that omitted extreme outliers. Controls are included for the poverty level of the school and for relative levels of implementation fidelity. A multiple base-line interrupted time series analysis, which is especially well suited to the analysis of aggregated

data from multiple groups, with varying characteristics and over a relatively long time span, supplements simple t-test comparisons.

The results are similar across all analyses, with significantly higher achievement scores in school years in which Direct Instruction was implemented. Table 14 summarizes some of the results. The first two columns of data summarize results with the analysis of first grade achievement, and the remaining columns summarize results with the analyses of fifth grade achievement, with the final column summarizing the results with the cohort analysis. The top panel gives the effect sizes reported above. All of the effect sizes easily surpass the criterion of .25 usually used to denote educational importance. They range from .27 for the cohort analysis to .86 for the analysis of fifth grade achievement with the full sample. The average effect size across the comparisons is .43.

Table 14
Summary of Effect Sizes Across All Analyses

	First Grade		Fifth Grade		
	Full Sample	Reduced Sample	Full Sample	Reduced Sample	Cohort Analysis
Effect Sizes					
Total Group	0.29	0.29	0.47	0.38	-----
Low Poverty Sch. Omitted	0.40	0.38	0.61	0.39	-----
From Mixed Model	0.39	0.44	0.86	0.43	0.27
Between School Variation Related Statistics					
Correlation Ratio	0.12	0.14	0.28	0.37	0.52
PRE Measure	0.19	0.11	0.72	0.59	0.79

Note: The reduced sample used in the analysis of first and fifth grade omitted outliers, school years with scores above the 90th and below the 10th percentile. Details on the calculations of the effect size from the mixed models, the correlation ratios, and the PRE measure are in the text and associated footnotes.

In short, SAT reading scores for first and fifth graders in GPSS were higher in schools and years when DI was implemented. These differences are both statistically significant and educationally important. They appear both when controls are introduced for the poverty status of the schools and when they are not included (compare the first row of data with the second and third rows in Table 14). They appear with analyses of both first and fifth grade achievement and when the sample is limited to cohorts that had data for both grade levels.

The bottom two rows of data in Table 14 summarize statistics related to the variation between schools in achievement. The correlation ratio indicates the proportion of total

variation that is between schools. The values are more than twice as large for the fifth grade analysis as in the first grade analysis and even greater for the cohort analysis. In other words, at the higher grade more of the variation in the average scores of the schools is between schools rather than over time. The results associated with the PRE measure are even more striking, for this statistic indicates the extent to which the variables in the mixed model (having DI and the rate of school poverty) can explain the differences between the schools. While these values are less than .20 for first grade they are substantial for the fifth grade analyses, ranging from .59 to .79. In other words, approximately three fifths or more of the differences between schools in average fifth grade achievement over time can be explained by the poverty level of the schools and by whether or not they had implemented Direct Instruction, suggesting that the impact of having Direct Instruction was even greater for the older students.

The results of this analysis replicate many other studies that have found that students exposed to Direct Instruction have higher achievement scores, and that this impact occurs throughout the elementary years. The strong effects on achievement in fifth grade are especially striking. It should be recalled that neither higher first grade achievement nor having DI in first grade was sufficient to produce higher average fifth grade scores. Instead, exposure to the program in fifth grade had a strong and significant impact on average achievement. Also notable is the extent to which differential exposure to the program can account for variations between the schools in fifth grade achievement.

One troubling aspect of the results should not be ignored – the decline in average scores from first to fifth grade. Inspection of the results from analyses of the total group and for the sub-group involved in the cohort analysis indicates a substantial decline in average scores from early to late elementary school. While this decline appears in all schools, it is significantly smaller in those that had implemented DI in fifth grade. It is also possible that the extent of the decline is exaggerated by the use of aggregate, rather than individual level, data; and that, if the analysis had tracked the progress of individual students, alternative results could have appeared. Nevertheless, such a decline is troubling and is far from inevitable. A recent article examined the impact of continuing exposure to Direct Instruction on student achievement in a very high poverty, inner-city setting. The results, using individual-level data indicate that those with exposure to the program in both first and fifth grade had significantly greater gains in reading achievement over time than students in other settings and that their average levels of achievement in fifth grade were above the national average (Stockard, 2010).

Several limitations to this report should be noted. Most important is the fact that only aggregate data were available for analysis, albeit for a relatively long period of time. Individual level data that would allow tracking of students through their school careers

would, of course, be ideal to examine the impact of the intervention.[8] Measures of implementation fidelity were global in nature and not specific to grade levels within schools or to the years of implementation. There was also no independent confirmation of the validity of the testing procedures that were used. No data were available for 2008 and data from 2009 to 2011 were manually extracted from power point slides of an administrator's public presentation. Having actual copies of the test results would, obviously, have been preferable.

Nevertheless, the analysis provides a potentially important test of a large implementation of Direct Instruction. Anecdotal evidence indicates that there were many challenges to the implementation involving administrative issues at the district and school levels and issues involving district-level policy makers. Despite the many challenges, the data from both first and fifth grade indicate that students exposed to Direct Instruction had higher levels of achievement and that these differences were statistically significant and educationally important.

[8] In contrast to other literature, the analyses reported here indicated no significant association of having DI in kindergarten and implementation fidelity with average levels of achievement. It could be hypothesized that such a relationship would appear if individual level data had been analyzed.

References

Benner, G.J., Nelson, J.R., Stage, S.A., & Ralston, N.C. (2011). The influence of fidelity of implementation on the reading outcomes of middle school students experiencing reading difficulties. *Remedial and Special Education, 32*, 79-88.

Biglan, A., Ary, D., & Wagenaar, A.C. (2000). The value of interrupted time-series experiments for community intervention research. *Prevention Science, 1*, 31-49.

Biglan, A. Flay, B.R., Komro, K.A., Wagenaar, A.C., Kjellstrand, J. (2012). Adaptive time-series designs for evaluating complex multicomponent interventions in neighborhoods and communities, Oregon Research Institute.

Carlson, C. D., & Francis, D.J. (2002). Increasing the reading achievement of at-risk children through Direct Instruction: Evaluation of the Rodeo Institute for Teacher Excellence (RITE). *Journal of Education for Students Placed at Risk, 7* (2), 141-166.

Gersten, R.M., Carnine, D.W., & Williams, P.B. (1982). Measuring implementation of a structured educational model in an urban school district: An observational approach. *Educational Evaluation and Policy Analysis 4(1)*, 67-79.

Gersten, R., Carnine, D., Zoref, L. & Cronin, D. (1986). A multifaceted study of change in seven inner-city schools. *The Elementary School Journal 86* (3), 257-276.

Madigan, K. & Cross, R.W. (2012). Impact of a schoolwide positive behavioral and intervention supports model on academic achievement in K-12 grades. Accountability Works, Inc.

Pearson Publishing Company. (2003). *Stanford Achievement Test Series, Tenth Edition, Spring Multilevel Norms Book*.

Robinson, L., Towner, J., Caros, J. & Billups, L. (2008). *Final Report: External Evaluation of Guam Public School System (GPSS) Direct Instruction Intervention Implementation, 2003-2008.* School Evaluation and Instructional Research Associates

Ross, S.M., Nummery, J.A., Goldfeder, E., McDonald, A., Racho, R., Hornbeck, M., & Fleishman, S. (2004). Using school reform models to improve reading achievement: A longitudinal study of Direct Instruction and Success for All in an urban district. *Journal of Education for Students Placed at Risk, 9*, 357-368.

Stockard, J. (2010). Promoting reading achievement and countering the "fourth-grade slump": The impact of Direct Instruction on reading achievement in fifth grade. *Journal of Education for Students Placed at Risk, 15,* 218-240.

Stockard, J. (2011). Direct Instruction and first grade reading achievement: The role of technical support and time of implementation, *Journal of Direct Instruction*, 11, 31-50.g

Stockard, J. (2013). Direct Instruction in the Guam Public Schools: An Analysis of Changes in Stanford Achievement Test Scores, Technical Report 2013-2.

Tallmadge, G. K. (1977). *The Joint Dissemination Review Panel Ideabook.* Washington DC: National Institute of Education and U.S. Office of Education.