

Response to Reading First Interim Report
Dr. James A. Salzman, Co-Director
Reading First Ohio Center

To paraphrase Mark Twain: There are lies, damned lies, and the latest [Reading First report](#). The report is methodologically flawed, statistically glamorous, and ultimately meaningless in terms of its conclusions. It's caused the usual sharks to roil the waters as if chum were being served. And in the end, it says nothing about the positive impact of Reading First in Ohio.

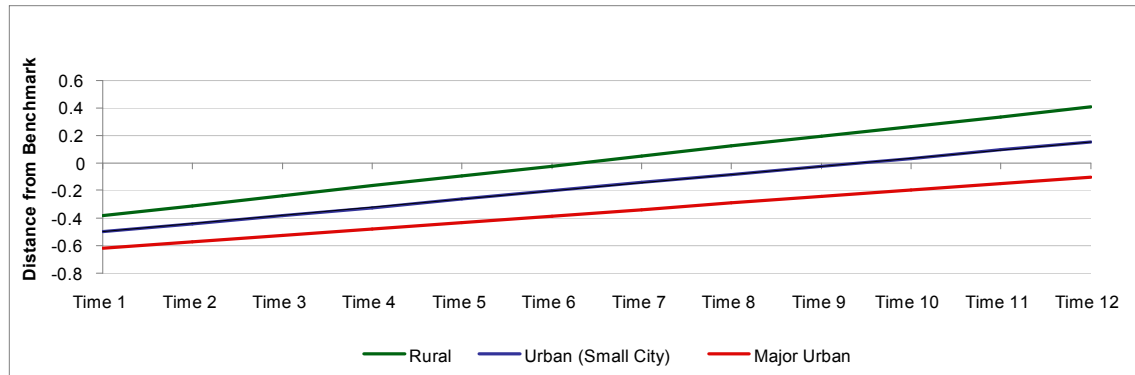
Flawed Methodology

Let's talk methodology first. One of the first things that any educator learns in a research class is that there are standards to follow in determining how large a sample you need to generalize to a population. The Institute of Education Sciences (IES) wants its report to be read as authoritative, though they recognize its *preliminary* nature. Therefore, they are asking readers to accept that their sample size was chosen in such a way as to make the results generalizable to all Reading First schools. Their sample consisted of 18 sites with 128 Reading First schools. In 2006, there were approximately 1,550 districts with 5,200 schools in Reading First. In order to have a 95% confidence level that their results could be generalized, a conservative estimate is that they would have needed about 886 schools in their sample. In its best light, the IES report may be able to be generalized to the impact of Reading First at the sites they used, but that's about it. That allows IES to discuss the impact on one state program and a total of 34 districts, or approximately 2% of Reading First districts. Given this serious flaw, the surprise is that groups, like the International Reading Association, who should know better and who certainly have not been shy about noting methodological concerns when it served their purposes, have not only not raised concerns but have jumped onto the pile.

The previous estimate of 886 schools assumes that the sample is randomly chosen. However, the sample for the IES study was not randomly chosen but was [purposive](#), or chosen so that the schools met a certain set of criteria. This is problematic for a number of reasons. For brevity's sake, though, allow me to focus on the major issue. The schools that are analyzed must come from districts that are large enough to fit the criteria that IES set to run their impressive-sounding statistical analysis. That means that most of the schools likely came from large urban districts. This ignores any impact that Reading First is making in rural and smaller districts; these are the very districts that we have seen the greatest growth in Reading achievement in Ohio (see Figure 1).

Figure 1

Linear Growth of Students in Reading First Schools on DIBELS Benchmark Tests from 2003-07 Disaggregated by School Type (Rural, Urban, and Major Urban



Note: The data in this chart come from all students (N=52,403) who have participated in Reading First in the schools in all three cohorts. For purposes of reading this graph, a flat line would indicate students having a year's growth in reading per year of instruction; a positive slope indicates more than a year's growth.

Frighteningly Flawed Thinking

There is another point to be made about the purposive sample. To briefly describe IES's methodology, the schools were chosen in the following way. If I am in a high-poverty, low-achieving district with 20 schools *and* I simply rank-ordered my need based on a quantitative measure, such as "past student reading performance or poverty" (IES, 2008, p. xi), then I would be eligible to be in the IES study. If I had enough funding for 10 schools and ranked my schools from 1 to 20 with 1 being the highest need and 20 the lowest, IES did their analyses on schools 6 through 10 (those in Reading First or RF) by comparing them to schools 11 through 15 (those who were next in line to receive funding). In their explanation for this methodology, the authors note that this method allows the researchers to control for "all systematic *pre-existing* [italics added] differences between the two groups" (IES, 2008, p. xi).

Since IES can control for pre-existing differences, the underlying conceptual assumption that provides the foundation for the study is that schools are either *doing Reading First* or not. IES would argue that, by seeing the difference in the ways that students achieve in RF schools and comparing them with what happens in closely matched schools that did not receive RF funds, they can draw a conclusion on the impact of the program. This is only a portion of the picture. While it is not clear that other urbans are acting in ways similar to those in Ohio, I can say with authority that in our large urban districts, the central office was not standing still while RF was being done in their district. Cleveland Municipal School District (CMSD) is a good example. While 20 schools were funded by Reading First in the district, the district spent their own funds to run a parallel program

that infused scientifically-based reading research (SBRR) strategies and practices alongside a SBRR-based core reading program. In essence, CMSD was running a Reading First-lite program, taking the professional development and resources from RF and leveraging them with district funds. Under this scenario, it's not unexpected that IES cannot tease out any differences among RF and non-RF schools and, in fact, the lack of differences may say more about districts in the study recognizing the importance of SBRR practices and use of a strong core program. If this is true, then it is further evidence of the *positive impact* of Reading First.

The previous points are admittedly speculative because, to be meaningful, the study would have to include the correct variables to analyze. I would argue that there are two variables that could have been measured and weren't. One variable would measure how well each of these schools implemented the expectations of RF. Another important variable would be something that indicated how much of the practices that define Reading First expectations (e.g., use of valid and reliable assessments, dedicated time for reading, SBRR strategies explicitly taught, etc.) are being implemented, and to what degree, in non-RF schools. Lacking this, all we know is which schools had money provided by Reading First. It should be clear to all by now that money alone will not make a difference unless teachers are doing the right things for students with that money. And, again, if other districts are doing work similar to Cleveland, they aren't waiting for the results of quasi-experimental studies to get them to move on something that makes sense to them.

The Bling of Glamorous Statistical Analyses

When a researcher or evaluator says *discontinuity regression analysis* or some other tongue-twisting data analytic technique, a number of readers probably just assume that this must mean that the results are just what the researcher claims they are. In this case, IES used a technique that has the potential to do exactly what they claim. However, they have put the statistical cart in front of the methodological horse. They have used a statistical technique, chosen a dataset that meets the assumptions of the technique, and in so doing, have lost any meaningful conclusions. I am hopeful that the previous discussion has shed some light on this issue.

Regression is a statistical technique that allows for the co-varying of multiple variables; the strength of the technique is that an evaluator can include important variables from the "real world" that they hypothesize may influence the success or failure of an initiative. This statistical technique, however, is not limited to controlling for pre-existing conditions (something that is very important in comparing two sets of schools) but can also control for any differences that exist *during* the time of the study. If IES had identified ways to measure other variables, like those mentioned above, they would have at least had a more defensible argument to make. Even if they had done this, though, it would not have overcome the serious flaws in the sample. In the end, IES had at their disposal the entire database of RF schools and states. That they chose to use a technique that compelled them to focus on about 2% of RF schools is either a stupendous

miscalculation or seems to prove the adage of their study design being [good enough for government work](#).

Reading First in Ohio

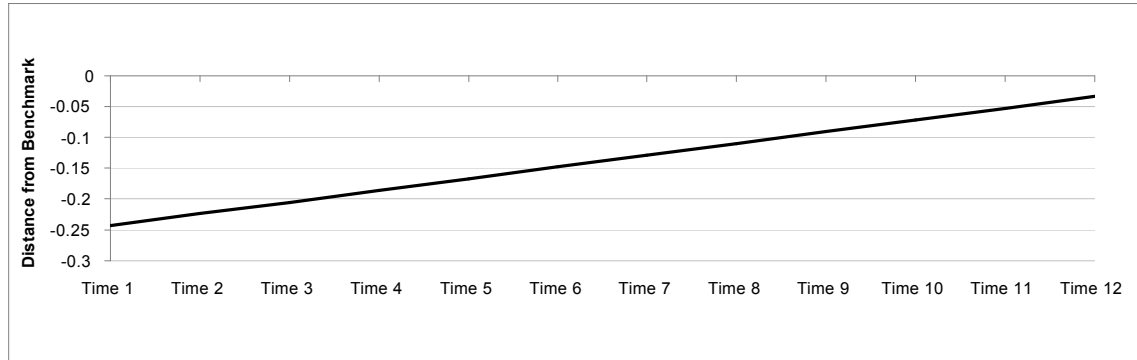
It is disappointing that there are now calls recommending an end to a program without fully reading and understanding the report that was just released. It is equally disappointing that there are very few education writers who have taken a critical look at the report; those who have come to many of the [same conclusions of the weakness](#) of the report that I have.

As educators in Ohio, then, it is important that you know the story of the success of Reading First in the state. Over the past five years, Reading First Ohio has helped teachers help more students in our highest poverty-lowest achieving schools read better. In addition to Figure 1 above, the data show that:

- Students in Ohio have gained more than a year's reading achievement for each year that they are in the program. This has held true for both Cohort 1 (see Figure 2) and Cohort 2 (see Figure 3). These figures also show that, if students stay within the program, they are able to catch up to benchmark scores in fluency, even though they start significantly behind.
- Students have closed the gap on state performance on the third grade Ohio Achievement Test (OAT) over the past four years (see Figure 4).
- Teachers have helped students close the achievement gap for students of color (see Figure 5; for more in depth look at this issue, see also Salzman, Newman, Clay, Brown, & Lenhart, 2007; Salzman, Newman, Clay, & Lenhart, 2006).
- Equally importantly, Westat's (2008) independent evaluation of Reading First Ohio has documented that the more time that students spend in Reading First schools the more they outperformed their peers in comparison schools across the state.

Figure 2

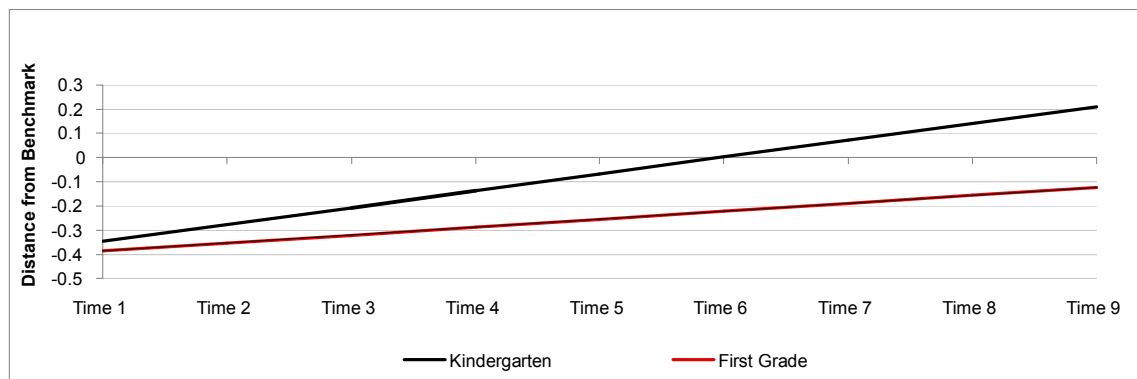
Linear Growth of Full-Treatment Students in Cohort 1 Reading First Schools on DIBELS Benchmark Tests from 2003-07



Note: The data in this chart come from Cohort 1 students (N=1,513) who began kindergarten in 2003 and participated in all Reading First testing through third grade in 2007. For purposes of reading this graph, a flat line would indicate students having a year's growth in reading per year of instruction; a positive slope indicates more than a year's growth.

Figure 3

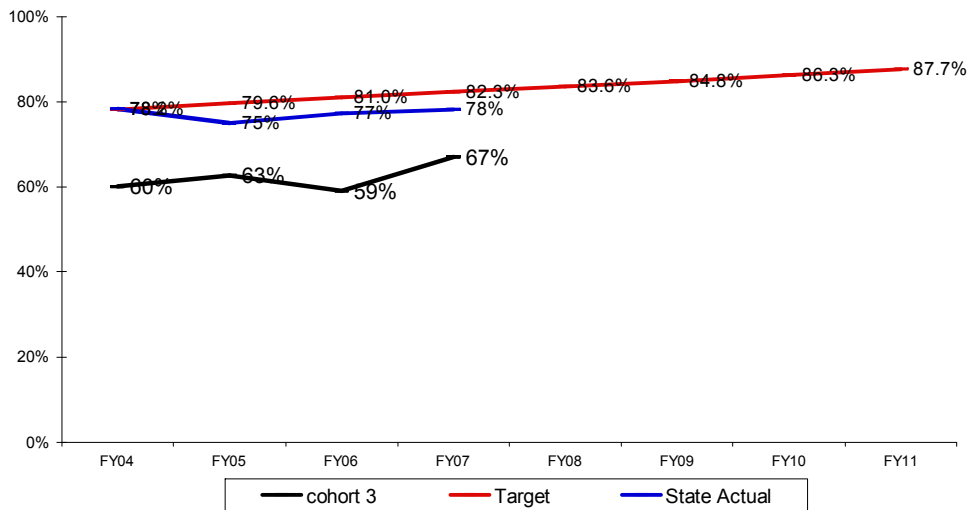
Linear Growth of Full-Treatment Students in Cohort 2 Reading First Schools on DIBELS Benchmark Tests from 2004-07



Note: The data in this chart come from Cohort 2 students who began kindergarten (N=711) and first grade (N=802) in 2004 and participated in all Reading First testing through second and third grade, respectively, in 2007. For purposes of reading this graph, a flat line would indicate students having a year's growth in reading per year of instruction; a positive slope indicates more than a year's growth.

Figure 4

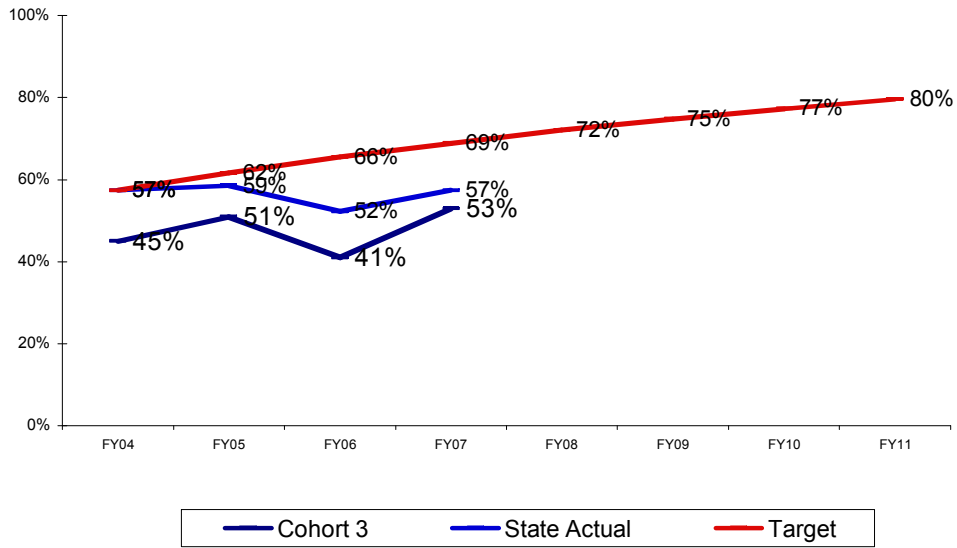
Comparison of Cohort 3 Reading First Schools, State Mean Scores, and Statewide Targets on 3rd Grade Students' Performance on Ohio Achievement Test (OAT) in Reading from 2004-07



Note: Students (N=1908) in the RFOC database are third graders in Cohort 3 Reading First Schools (N=47) who began the program in 2004; State Actual represents all test takers with data downloaded from ODE's website.

Figure 5

Comparison of Cohort 3 Reading First Schools, State Mean Scores, and Statewide Targets on 3rd Grade African-American Students' Performance on Ohio Achievement Test (OAT) in Reading from 2004-07



Note: Students (N=1908) in the RFOC database are third graders in Cohort 3 Reading First Schools (N=47) who began the program in 2004; State Actual represents all test takers with data downloaded from ODE's website.