

RESPONSE TO THE READING FIRST IMPACT STUDY INTERIM REPORT

20 August 2008

Reading First Federal Advisory Committee Members

Katherine Mitchell (Chair)
Former Assistant State Superintendent of Education for Reading,
Alabama Department of Education

Susan Brady
Haskins Laboratories (Yale), and Department of Psychology,
University of Rhode Island

Laurie Cutting
Kennedy Krieger Institute and
Johns Hopkins University

Maria Carlo
School of Education,
University of Miami

Kris Gutierrez
Graduate School of Education and Information Studies,
University of California, Los Angeles

Nonie K. Lesaux
Graduate School of Education,
Harvard University

John Wills Lloyd
Curry School of Education,
University of Virginia

Frederick Morrison
Center for Human Growth and Development,
University of Michigan

Annemarie Sullivan Palincsar
School of Education,
University of Michigan

Donna Scanlon
School of Education, University at Albany,
State University of New York

Frank Vellutino
School of Education, University at Albany,
State University of New York

Contents

Executive Summary	i
The Reading First Program	1
The Reading First Impact Study	1
Major Limitations: Lack of Random Assignment and Contamination of Comparison Condition	3
Additional Limitations in the Design and Implementation of the Study	4
Unrepresentative Sample.....	4
Limited Variables Assessed.....	5
What Cannot Be Concluded from the Reading First Impact Study.....	6
Closing Comments and Recommendations.....	7
Other Sources of Data	7
Rigorous Evaluations	8
Recommendations.....	9
References	10
Appendices	11
A. Reading First Purposes.....	11
B. Authorized Analyses and Measurements.....	12

Executive Summary

We, the members of Reading First Federal Advisory Committee, recommend caution in interpreting results from the *Reading First Impact Study: Interim Report*. Although the Impact Study has technical strengths, it has fundamental flaws that severely limit its utility for informing policy decisions. In this document we present the reasons for our reservations about the study and our recommendations for further action.

Major Limitations: Lack of Random Assignment and Contamination of Comparison Condition

We have important reservations about the design and implementation of the study. The most critical concern has to do with the basic design of the study. Although the Reading First legislation called for a study in which schools would be randomly assigned to Reading First and non-Reading First treatments, random assignment was not employed. Instead, the evaluation only included schools that were in districts that had received Reading First funding. Within those local education agencies, comparisons were made between schools that were awarded Reading First grants and those that were not awarded Reading First grants. Districts made the decision regarding which schools would receive funding and this was not done randomly. In addition, the evaluators failed to consider the impact that Reading First had at the district level. Most districts that received Reading First funding used the funding to build district-wide capacity to implement the scientifically-based instructional and assessment practices that are encouraged by Reading First. As a result, schools that did not receive Reading First grants often received the same professional development opportunities and adopted the same instructional materials as schools that received Reading First grants. Because of this contamination across conditions, the only reliable difference between the two groups of schools in the Impact Study was the source of funds that supported instructional changes. Thus, the study simply does not assess the effects of implementing the instructional practices encouraged by Reading First; therefore, it does not provide an evaluation of the impact or effectiveness of Reading First.

Additional Limitations in the Design and Implementation of the Study

Even if the evaluation did not have the substantial limitations we identified in the previous section, it would still have other limitations that would make us reluctant to endorse its findings for purposes of making policy decisions. In this section we discuss problems associated with the sample assessed and with the set of outcomes examined.

Unrepresentative Sample

The schools studied in the Reading First Impact Study were not representative of schools receiving funding under Reading First. The sample of schools in the study was limited to those that had been identified for funding according to a rank-ordering procedure. In that procedure, local or state education agencies ranked applicants on variables that indicated their need and then

used a cut point to determine which schools would receive Reading First funding: Schools below the cut point received funds. The evaluators identified 1 state and 29 (of 1,890) local education agencies that employed this method of awarding Reading First funds. Of the 29 local education agencies that used this procedure, a subset of 17 actually participated in the study (Gamse, Bloom, Kemple, & Jaccob, 2008). Thus, the Reading First Impact Study is based only on a sample of approximately 2% of the local education agencies and schools that received funding under Reading First. Differences between the schools used in the Impact Study and the much larger population of schools funded under Reading First precludes generalizing any outcomes assessed by the study.

Limited Variables Assessed

Although the Interim Report of the Reading First Impact Study analyzed results for reading comprehension and a few classroom practice variables, the range of variables assessed does not provide a thorough or adequate analysis of the outcomes of Reading First. Without examining effects on other key literacy-related outcomes (e.g., phoneme awareness, reading fluency, vocabulary knowledge), it is not possible to determine whether Reading First had beneficial effects. Given that research-based instructional practices regularly have been found to affect component skills that influence reading comprehension, it is important to consider evidence of effects on other aspects of reading competence in addition to comprehension. Although the final report of the Impact Study will include some additional measures, the scope still will not be sufficient to evaluate the effects of Reading First funding adequately, nor will it overcome the fundamental problem of diffusion of treatment discussed in the Major Limitations section summary.

In addition, the Interim Report of the Impact Study does not provide information about key process variables (e.g., curricula employed, quality of professional development), with the consequence that it cannot inform the American public regarding the relative contributions of these variables to student success. Members of the Reading First Federal Advisory Committee do not expect that all states, local education agencies, schools, or teachers will have equal outcomes, so we and others concerned with improving educational practice need to know under what conditions greater success occurs.

What Cannot Be Concluded from the Reading First Impact Study

Critics of Reading First have portrayed the Reading First Impact Study as evidence that Reading First “doesn’t work.” As we have sketched in the foregoing paragraphs and detailed in the full document, we consider the evidence from the Reading First Impact Study inadequate to make such judgments. It did not compare Reading First schools to schools not using practices mandated under Reading First. It only studied a small, unrepresentative sample of schools. It did not provide data on other relevant outcomes and potential moderator variables.

Likewise, contrary to the claims of critics, it is important to note that the Reading First Impact Study was not designed to evaluate the efficacy of scientifically-based reading assessment and instruction. The value of teaching the five essential components of reading specified in the Reading First legislation and the efficacy of instructional approaches designed to develop those domains have been validated in numerous carefully-controlled reading research studies, as documented in the National Reading Panel report (National Institute of Child Health and Human Development, 2000) as well as extensive, more-recent, scientifically rigorous research. This growing body of research continues to substantiate the rationale for Reading First.

Closing Comments and Recommendations

A final area of concern is the focus on the *Reading First Impact Study Interim Report* in the press and by members of Congress without consideration of other sources of information about Reading First. The Reading First Advisory Committee strongly urges members of Congress to make decisions about the future of Reading First only when the multiple sources of information available have been reviewed and analyzed. These sources include the final report of the Reading First Impact Study (with the caution that it will not be able to correct the problems that we have described), the annual Reading First progress reports required of each state and territory, and the independent evaluation studies being conducted in each state and territory. Those sources should also include results from the comprehensive list of analyses outlined in section 1205 of the Reading First legislation.

In offering this critique, the committee's intent is not to criticize the evaluation team. We have identified shortcomings in the evaluation because of the potential influence the report may have on Reading First policy. The contractors employed excellent technical strategies and sophisticated analytic procedures, creating a methodologically rigorous evaluation that is one of the very few comprehensive, empirical studies of a national policy. The problems we cite arose because of constraints over which the evaluators had virtually no control.

Based on the concerns expressed in this statement, however, the Reading First Advisory Committee respectfully makes the following recommendations to members of Congress:

1. Refrain from eliminating funding for Reading First.
2. Regard the final report of the Reading First Impact Study with caution, because it will not be able to correct the fundamental problems we have described.
3. Defer further action regarding the future of Reading First until data from the multiple sources available have been reviewed and analyzed.
4. Incorporate funding and authority for rigorous evaluations in future legislation, extending and improving the ground breaking plans that were part of the Reading First legislation.

RESPONSE TO THE READING FIRST IMPACT STUDY INTERIM REPORT

This statement extends and amplifies a previous statement developed by the Reading First Advisory Committee following a meeting of the committee held 23 June 2008, to discuss findings summarized in the *Reading First Impact Study: Interim Report*. During our discussion of this report, the committee identified three key concerns associated with results summarized in the interim report and interpretation of those results:

1. Limitations surrounding the design, methods, and analyses of the Reading First Impact Study;
2. Unwarranted inferences and conclusions being drawn from the Reading First Impact Study Interim Report;
3. Failure to consider other sources of information about the effectiveness of Reading First.

In the initial statement, the committee promised to generate a brief document describing the limitations of the study and explaining what cannot be concluded from the interim report. This document fulfills that commitment, but first we begin with brief summaries of both the Reading First program and the Reading First Impact Study as described in the interim report.

The Reading First Program

The Reading First program is a large-scale intervention program designed to improve reading achievement in struggling readers being educated in high-poverty, low-performing schools. Funding for this program was authorized by Congress as a part of the No Child Left Behind Act in 2001 for purposes of supporting: (1) reading curricula and materials that focus on five of the major components of reading (phoneme awareness, phonics, vocabulary, reading fluency, and reading comprehension); (2) professional development and coaching of teachers in the use of scientifically-based reading practices in both typical classroom and remedial settings; and (3) identification and prevention of early reading difficulties through student screening, intervention for struggling readers, and ongoing monitoring of students' progress. (See Appendix A for a formal list of Reading First's purposes). Reading First grants were awarded to state education agencies between July 2002 and October 2003. By April 2007, states had awarded Reading First sub-grants to 1,809 school districts that, in turn, provided funds for 5,880 schools within those districts (Gamse et al., 2008).

The Reading First Impact Study

Congress authorized funds for and mandated a comprehensive evaluation of the impact and effectiveness of the Reading First program. In addition to an overall analysis of the effectiveness of the program, the legislation called for study of the relationships among the components of reading instruction and reading proficiency, adequacy of implementation at the state and local

levels, impact of specific reading programs and assessment procedures, readiness of future teachers to provide research-based methods of instruction, and more (see Appendix B for a list of the analyses that Congress envisioned for this evaluation).

The Institute for Education Sciences commissioned the Reading First Impact Study to evaluate the impact of the Reading First program. Contracts for the study were awarded to ABT Associates as primary contractors and to MDRC, Westat, and others as subcontractors or collaborators. The study was conducted over the course of three school years (2004-2007) to address the following questions: (1) What is the impact of Reading First on student reading achievement? (2) What is the impact of Reading First on classroom instruction? (3) What is the relationship between the degree of implementation of scientifically based reading instruction and student reading achievement? The interim report presents results from data collected during the first two years of the study and provides preliminary responses to aspects of the first two questions. A forthcoming final report, which will include the third year of data, is intended to address all three questions as much as possible, given the constraints imposed on the project.

The initiation of the Reading First Impact Study was significantly delayed and many school districts that received Reading First awards had begun implementation of mandated instructional practices before the Impact Study began. As a result, in the districts that were already funded, it was not possible to assign eligible schools randomly to treatment and control conditions. Because random assignment was not possible, the evaluators elected to use a regression discontinuity design for the study. In the context of the Reading First Impact Study, this design required the use of districts in which schools had been assigned to the Reading First and non-Reading First conditions on the basis of their status on rating scales developed by their districts (e.g., past reading performance, poverty level). Because only 29 of the 1,890 districts that received Reading First funding used this sort of procedure for making funding decisions, the design greatly reduced the number of school districts eligible for the study. From the 29 potential sites, 17 were ultimately selected for inclusion. Thus, out of 1,890 school districts and 5,880 schools that received Reading First funds, the Impact Study was conducted in only 17 school districts across 12 states and in one statewide program (Gamse et al., 2008).

The final analysis included 119 schools that were funded through Reading First and 119 schools that were not. Outcome measures assessing the impact of the Reading First program included a standardized test of reading comprehension (assessed in grades 1, 2, and 3), classroom observations of the amount of time teachers spent on five major components of reading and reading instruction (assessed in grades 1 and 2), and classroom observations of student engagement with print (assessed in grades 1 and 2). Analyses were conducted that included all study sites and additional analyses were conducted for study sites that received their Reading First grants somewhat earlier in the federal funding process (between April and December 2003) and those that received grants later in the federal funding process (between January and August 2004; Gamse et al., 2008).

Findings from the Reading First Impact Study were mixed. When averaged across all 18 sites, the estimated difference between funded and non-funded schools on the test of reading comprehension was not statistically significant. However, in the separate analyses for so-called early and late awards sites, a different pattern of results emerged. In late award sites, funded schools performed better than non-funded schools on the measure of reading comprehension at all three grade levels and some of these differences were statistically significant (see page xvii of the Interim Report Executive Summary).

Another finding of the study was that, across all 18 sites, schools with Reading First grants devoted significantly more total class time to the five essential components of reading instruction promoted by the program. Separate analyses of the early and late award sites revealed that these differences were most apparent in the late award sites and not significant in the early award sites. In the overall analysis, it was evident that funded schools allotted more time for explicit instruction in grades one and two, and for quality student practice in grade two. The program had no (apparent) positive effect on student engagement with print and average impacts on reading comprehension and classroom instruction did not change systematically over time, as sites gained experience with the Reading First program.

Major Limitations: Lack of Random Assignment and Contamination of Comparison Condition

The Reading First legislation specified a study in which schools would be randomly assigned to Reading First and non-Reading First treatments. However, random assignment was not employed in the Reading First Impact Study. As noted earlier, because the study was implemented late, it was not possible to assign schools randomly to the Reading First and Non-Reading First conditions. Across the nation, most states had begun implementing Reading First. Consequently, eligible schools had already been placed in one or the other condition and this had not been done randomly. In lieu of the optimal randomized control study, the Reading First Impact Study used a regression discontinuity design that is generally considered to be a good alternative to a randomized control design. To be eligible for inclusion in a study using this design, districts had to have used a rank ordering procedure to determine which schools received Reading First Funding and which schools did not receive such funding. Those above the district-designated cut scores did not receive funding, and those below the cut scores did receive funding. Schools included in the study are those that scored a little above the cut score and those that scored a little below. The purpose of selecting schools in this way was to ensure that the schools in the treatment group (Reading First funded schools) and those in the comparison group (schools not funded) were as similar as possible to one another at the outset of the study. The basic logic of the design hinges on the assumption that only the funded schools would be affected by the Reading First program and, therefore, comparisons between funded and non-funded schools would provide evidence of the impact that Reading First had on students' reading performance and teachers' reading instruction.

However, the design fails to take into account the fact that states and districts that received Reading First funds were encouraged to use those funds for professional development in non-Reading First schools. As a result, in districts that received Reading First grants, the schools in the comparison group were allowed to participate in all Reading First professional development activities. In fact, many states and districts augmented the Reading First funds to ensure that the assessment and instructional programs were similar in schools that received Reading First grants and those that did not receive these grants. Indeed, at least 10 states (Alabama, Alaska, Arizona, California, Idaho, Kentucky, Maryland, Montana, Washington, and Wyoming) sent reports to the Reading First Advisory Committee indicating that this situation characterized their states. Because the schools that did not receive Reading First grants often received the same professional development opportunities and adopted the same assessment procedures and instructional materials as schools that did receive Reading First grants, it is logical that there would be few differences between the two groups of schools.

To evaluate the effects of a treatment (in this case Reading First funds), there needs to be a way to estimate how much of the treatment has been provided to the individuals or groups under study. Yet, because it was not possible to conduct a true randomized control study and because the evaluators failed to assess the degree of exposure and access that non-funded schools had to Reading First supported practices and materials, it is not possible to assess the impact of the treatment confidently. The only thing we know for sure about the two groups of schools compared in the Impact Study is that one group received funding directly from the Reading First program and the other did not receive such funding.

Additional Limitations in the Design and Implementation of the Study

Our most substantial concerns with the Impact Study are that there was not random assignment and that there was potential contamination of the comparison group. In addition, we have reservations about two other aspects of the Impact Study: (1) the study examined a sample of schools that appears unrepresentative of the larger group of schools in Reading First, thereby limiting generalizations that can be made, and (2) the set of variables assessed was limited, hampering evaluation of students' reading achievement and of factors related to the success of implementation. We elaborate on these issues in the following sections.

Unrepresentative Sample

The Reading First Advisory Committee is concerned that the Reading First Impact Study involved a group of Reading First schools that is not representative of schools in the states and territories that received funding under Reading First. Because only a limited number of districts used a single rating scale for making funding decisions, the number of districts that could be included in the regression discontinuity design was greatly reduced. As a result, the sample consisted of approximately 2% of the schools that received funding under Reading First. Moreover, although the sample of schools for the Impact Study shares many important

characteristics with the 5880 Reading First schools, it differs in several ways. For example, the research sample differs from the larger group in ethnic distribution of participants, in the proportions of schools from urban and rural settings, and in the average size of schools. Accordingly, these differences prevent us from concluding that the Reading First Impact Study sample represents the nation-wide Reading First program.

To add substance to this point, the majority of schools that were not included in the research sample are in states that won the earlier Reading First grants awarded. It is possible (and some argue highly probable) that the states and schools that received the first awards also were those with strong programmatic features, infrastructure, and implementation plans already in place. It could be the case that a larger and more representative sample of schools from states and territories receiving awards earlier in the process would have produced stronger effects than the small and less representative sample of schools included in the Reading First Impact Study. Some suggestive evidence to support this possibility is provided in the interim report, which indicates that the average amount of daily instructional time spent in the five components of reading was 59 minutes in funded Reading First schools and 50 minutes in non-funded schools. Although this difference is statistically significant, neither figure comports with the 90 minutes of daily instruction coupled with the 20-40 minutes of daily supplemental instruction observed in and reported by many Reading First schools. This fact alone suggests that the reading programs in both sets of schools included in the Impact Study sample were something less than adequate.

Limited Variables Assessed

An additional limitation of the Reading First Impact Study is that most of the evaluative analyses described in the legislation (see Appendix B) were not performed in the study. A comprehensive evaluation of the impact of the Reading First program would have included assessments, not only of reading comprehension, but also of the other foundational components of reading acquisition and reading instruction (phoneme awareness, phonics, vocabulary, and reading fluency). To be sure, among the reading performance outcomes for children, reading comprehension is the foremost target, but it is also the most difficult variable to influence and to measure. Given that scientifically based reading instructional practices repeatedly have been found to affect component skills that influence reading comprehension, it is important to consider evidence of effects on other aspects of reading competence in addition to comprehension. At present, we do not know the extent to which strengths or weaknesses in each of the other four components of reading and reading instruction influenced reading comprehension outcomes.

A comprehensive evaluation of the Reading First program likewise would have included measures of the nature and quality of assessment and instruction in all five components of reading instruction. Had data from such measures been collected, it may have been possible to identify practices that were associated with stronger and weaker student outcomes and the

contamination problem noted previously would not have been quite so limiting. As pointed out by a noted methodologist in literacy research (David Francis, in a paper presented at the Institute for Education Sciences Research Conference, 11 June 2008), the Reading First Impact Study was not designed to estimate variability in treatment effects accurately or the effects of key moderators of treatment impacts. Because of the high degree of heterogeneity in implementation of key program components that likely characterized both Reading First and non-Reading First schools (e.g., professional development, the degree to which assessment informed instruction, educational leadership, percentage of struggling readers), it would be important to ascertain which factors moderated the effectiveness of treatments. This was not done. In alignment with this concern, consider the fact that the Reading First Impact Study provided no data that could explain the finding that later award sites produced consistently positive results on reading comprehension, whereas the earlier award sites in the study produced consistently negative results on this measure.

Members of the Reading First Federal Advisory Committee do not expect that all states, local education agencies, schools, or teachers will have equal outcomes, so we and others concerned with improving educational practice need to know under what conditions relatively greater success occurs. The Interim Report of the Reading First Impact Study does not address this issue. The consequence is that it cannot inform the American public regarding what worked and what did not work. The incomplete assessment of student achievement, reading programs, and implementation leave many crucial questions unanswered. Ultimately, we need to have answers to such questions as these: What features or aspects of given curricula produce stronger outcomes than others? Do some grouping practices enhance students' reading performance? Are certain features of professional development associated with improved student learning?

What Cannot Be Concluded from the Reading First Impact Study

Critics of Reading First have portrayed the Reading First Impact Study as evidence that Reading First “doesn’t work.” However, as asserted in this critique, we consider the evidence from the Impact Study inadequate to make such judgments. The use of non-funded schools from the same districts as funded schools raises the probable risk of dissemination of Reading First concepts and practices to those schools, with the likely outcome that differences between teacher practices and student achievement in the two conditions were substantially reduced. Further, only a very small proportion of Reading First Schools were studied, and the sample was not representative in terms of demographic and achievement characteristics of the full sample of Reading First schools. In addition, the Reading First Impact Study did not provide data on outcomes for several relevant literacy skills targeted in Reading First, nor did it evaluate potential moderator variables such as the quality of professional development, the extent of support from coaches, or the curricula followed. In short, the study cannot evaluate whether, or under what conditions, Reading First “works.”

It also is important to note that the Reading First Impact Study was not intended to evaluate the efficacy of scientifically-based reading assessment and instruction: “This evaluation was not designed to be, nor is it, an evaluation of the effectiveness of scientifically-based reading instruction” (Grover R. Whitehurst, meeting with the Reading First Advisory Committee, 23 June 2008). The value of teaching the five essential components of reading specified in the Reading First legislation and the efficacy of instructional approaches designed to develop those domains have been validated in numerous carefully-controlled reading research studies, as documented in the National Reading Panel Report (National Institute of Child Health and Human Development 2000) and through extensive, more recent, scientifically-rigorous research.

Reading First does not constitute an evaluation of the efficacy of scientifically-based methods of reading instruction, but instead is an implementation project designed to disseminate what has been learned through research to improve the reading achievement of students in high-poverty, low-performing schools. What could be learned in studying Reading First is how effective particular dissemination techniques and procedures have been. The current study does not do this. The Reading First Impact Study might be thought of as an effectiveness study asking whether the receipt of Reading First funds improved the reading achievement of students. However, as previously noted in some detail, it seems likely that all of the schools in the study were affected by the Reading First program, at least to some extent, either through direct grants or through the services and financial supports offered by their districts or states or both. To an extent, the spread of effects to non-funded schools (which in our critique of the study we call a contamination problem) could be construed as a rather remarkable achievement of the Reading First legislation: it does appear that the federally provided funds were effectively leveraged by states and school districts in efforts to more broadly impact early literacy learners attending their high needs schools. Although this spread of effects caused major problems in efforts to assess the impact of the Reading First program on student achievement, it does speak in some ways to the success of the dissemination efforts.

Closing Comments and Recommendations

The previous points of discussion in this statement have pertained specifically to the Reading First Impact Study itself. In this section we explain two further concerns: The need to (a) consider other sources of information about Reading First outcomes and (b) promote rigorous evaluations of Reading First and other education-related policies.

Other Sources of Data

The Reading First Advisory Committee is concerned that congressional deliberations and actions regarding the future of the Reading First program may have been unduly influenced by results from the Reading First Impact Study interim report. We strongly urge members of Congress to defer further action regarding the future of Reading First until data from the multiple sources available have been reviewed and analyzed.

One such source is the forthcoming final report of the Reading First Impact Study. However, this document also must be viewed with caution, because it will not be able to correct the fundamental problems we have described, such as spread of Reading First influence throughout the districts studied and the lack of true, non-Reading First comparison schools.

Other important sources of information include the annual Reading First progress reports required of each state. In these reports, states present percentages of students in each grade who met the proficiency cut-off score set by that state on measures of reading proficiency. Because states were allowed to choose their own assessment instruments and set their own proficiency standards, the data cannot be easily used to compare results across states, but the data do make clear that in some states there was dramatic improvement over time in the number of children who scored at or above the set proficiency criterion. Although such increases cannot necessarily be directly attributed to the effects of Reading First, the fact that the changes coincide with the availability of Reading First funding certainly suggests that a relationship exists. Investigations of the procedures and materials adopted as part of the Reading First implementation in these more successful states is clearly warranted. The American Institute for Research has collected and summarized results from these annual performance reports and posted them on the Reading First Web site (<http://www.ed.gov/programs/readingfirst/>).

Another data source is the independent evaluations that have been and still are being conducted in each state. Reading First requires that each state conduct an independent annual evaluation of the effectiveness of the Reading First programs in schools receiving Reading First funds. Although the results again cannot be compared across all states, many of the states show gains in student achievement, helping explain why so many state officials claim effectiveness of the Reading First programs in their schools. Results from these evaluations also are available on the Reading First Web site.

Rigorous Evaluations

In closing, we urge government leaders and others to join us in commending federal educational programs that require outcome evaluations, a step that should be heralded, not abandoned. The gap between the degree and scope of evaluation requested by Congress initially and the degree and scope of evaluation attained in part reflects compromises that were made to make it easier for states to begin to participate in the evaluation.

The committee's purpose in this analysis is not to criticize the evaluation team. We have identified shortcomings in the evaluation, because of the potential influence the report may have on Reading First policy. The contractors employed excellent technical strategies and sophisticated analytic procedures, creating a methodologically rigorous evaluation that is one of the very few comprehensive, empirical studies of a national policy. The problems we cite arose because of constraints over which the evaluators had virtually no control.

One challenge for the future then is to adopt improved evaluation procedures and assist states in doing the same. Moving forward, we need to conduct the studies originally described in the Reading First legislation (see Appendix B). Those studies would allow us to learn more about the circumstances under which the approaches advocated in Reading First are most effectively implemented, on a broad scale, in schools serving our most challenged children. Likewise, it is critically important to use the information provided by the states' annual progress and independent evaluation reports, together with that from the Reading First Impact Study, to examine the conditions under which programs instituted with Reading First funding were or were not effective. The goals would be to ensure that the intervention program and its implementation can be strengthened and that our policies are achieving the results for which they were created. Ideally this would require an iterative process of rigorous evaluation and modification, which would be a laudable change from the usual funding of educational policies in the absence of any stringent information about their effectiveness. In contrast, abandoning support for educational practices that have science behind them would be a tragic step backwards.

Recommendations

In conclusion, the Reading First Federal Advisory Committee offers this statement to aid members of Congress and other policy-makers in making well-informed decisions about the future of Reading First. Based on the concerns expressed in this analysis, we respectfully make the following recommendations to members of Congress:

1. Refrain from eliminating funding for Reading First.
2. Regard the final report of the Reading First Impact Study with caution, because it will not be able to correct the fundamental problems we have described.
3. Defer further action regarding the future of Reading First until data from the multiple sources available have been reviewed and analyzed.
4. Incorporate funding and authority for rigorous evaluations in future legislation, extending and improving the ground breaking plans that were part of the Reading First legislation.

References

Gamse, B. C., Bloom, H. S., Kemple, J. J., Jacob, R. T. (2008). *Reading First Impact Study: Interim Report* (NCEE 2008-4016). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Regional Assistance, Institute of Education Sciences, U.S. Department of Education. National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Appendices

A. Reading First Purposes

This extract from Subpart 1 of Part B of the No Child Left Behind legislation identifies the purposes of Reading First

SEC. 1201. PURPOSES.

The purposes of this subpart are as follows:

(1) To provide assistance to State educational agencies and local educational agencies in establishing reading programs for students in kindergarten through grade 3 that are based on scientifically based reading research, to ensure that every student can read at grade level or above not later than the end of grade 3.

(2) To provide assistance to State educational agencies and local educational agencies in preparing teachers, including special education teachers, through professional development and other support, so the teachers can identify specific reading barriers facing their students and so the teachers have the tools to effectively help their students learn to read.

(3) To provide assistance to State educational agencies and local educational agencies in selecting or administering screening, diagnostic, and classroom-based instructional reading assessments.

(4) To provide assistance to State educational agencies and local educational agencies in selecting or developing effective instructional materials (including classroom-based materials to assist teachers in implementing the essential components of reading instruction), programs, learning systems, and strategies to implement methods that have been proven to prevent or remediate reading failure within a State.

(5) To strengthen coordination among schools, early literacy programs, and family literacy programs to improve reading achievement for all children.

B. Authorized Analyses and Measurements

This extract from Subpart 1 of Part B of the No Child Left Behind legislation identifies the goals of the evaluation of Reading First

SEC. 1205. EXTERNAL EVALUATION.

(a) **IN GENERAL.** From funds reserved under section 1202(b)(1)(C), the Secretary shall contract with an independent organization outside of the Department for a 5-year, rigorous, scientifically valid, quantitative evaluation of this subpart.

(b) **PROCESS.** The evaluation under subsection (a) shall be conducted by an organization that is capable of designing and carrying out an independent evaluation that identifies the effects of specific activities carried out by State educational agencies and local educational agencies under this subpart on improving reading instruction. Such evaluation shall take into account factors influencing student performance that are not controlled by teachers or education administrators.

(c) **ANALYSIS.** The evaluation under subsection (a) shall include the following:

(1) An analysis of the relationship between each of the essential components of reading instruction and overall reading proficiency.

(2) An analysis of whether assessment tools used by State educational agencies and local educational agencies measure the essential components of reading.

(3) An analysis of how State reading standards correlate with the essential components of reading instruction.

(4) An analysis of whether the receipt of a targeted assistance grant under section 1204 results in an increase in the number of children who read proficiently.

(5) A measurement of the extent to which specific instructional materials improve reading proficiency.

(6) A measurement of the extent to which specific screening, diagnostic, and classroom-based instructional reading assessments assist teachers in identifying specific reading deficiencies.

(7) A measurement of the extent to which professional development programs implemented by State educational agencies using funds received under this subpart improve reading instruction.

(8) A measurement of how well students preparing to enter the teaching profession are prepared to teach the essential components of reading instruction.

(9) An analysis of changes in students' interest in reading and time spent reading outside of school.

(10) Any other analysis or measurement pertinent to this subpart that is determined to be appropriate by the Secretary.