

Assessing the Rationales for Educational Reforms: An Examination of Policy Claims About Professional Development, Comprehensive Reform, and Direct Instruction

Edward P. St. John
Genevieve Manset-Williamson
Choong-Geun Chung
Robert S. Michael

Evaluation studies seldom examine whether the rationales used to argue for an educational reform actually hold up when empirical evidence is examined after the reform has been implemented. This article examines survey data from 3 years of analyses of early reading interventions to examine three rationales that were used to argue for the program. First, there was evidence to support the argument that teachers need time to collaborate about improving educational outcomes. This study found that 2 years of funding for early reading reforms provided a margin of difference for collaborative efforts among teachers to promote reading-related outcomes. Second, the argument that comprehensive reform strategies promote gains in student outcomes was also supported, but not all reform models had their intended effects. Finally, there was no evidence from this study that the direct/explicit approach to reading instruction improved student outcomes, although this claim merits more systematic study in the future.

Keywords: comprehensive school reform; state policy; reading programs; evaluation

Educational reform in the United States is at an interesting but perplexing point in its evolution. The current wave of school reform is, at least in theory, research based, which provides an opportunity for educational researchers

Authors' Note: An earlier version of this article was released by the Indiana Education Policy Center (Policy Research Report #01-03). The research reported herein was completed with the support of the Indiana Department of Education, state of Indiana. The interpretations in this article are the authors' and do not necessarily represent policies, positions, or views of the Indiana Department of Education any other agency of the state of Indiana.

DOI: 10.1177/0013161X04269618

© 2005 The University Council for Educational Administration

and perhaps even educators to test new approaches. Schools and school districts face choices about the types of reform strategies they may want to undertake. However, the long history of failure of education reforms, especially in urban schools (Miron & St. John, 2003), complicates efforts to develop policies and strategies to guide the choices educators might make about reform strategies. Different rationales have been used to argue for the new wave of reading and comprehensive reforms. This complicates efforts to evaluate the effect of education reforms.

Three arguments about reform seem to have a substantial influence on the way both recent reading and comprehensive reforms have been conceptualized within state and federal programs, implemented by states and schools, and evaluated. One is that professional development, especially providing opportunities for educators to collaborate in schools, can foster improvement in educational outcomes (Bull & Buechler, 1996; Education Commission of the States, 1997; Guskey & Sparks, 1997). A second is that comprehensive models that engage teachers in cohesive reform strategies increase the chances more students will succeed (e.g., Clay, 1993; Taylor, Anderson, Au, & Jaffy, 2000). A third is that direct, explicit instruction in letter/sound relationships is necessary to foster reading by third grade, a strategy that is essential to move more children into the educational mainstream (Foorman, Fletcher, Francis, & Schatschneider, 2000; Snow, Burns, & Griffin, 1998). All three rationales were used to argue for reading reform in Indiana. These same arguments also influenced reform strategies in other states.

Although these rationales are actually composed of micro sets of potentially competing reform arguments, each of the rationales is used in advocacy for research-based reform (i.e., Slavin, 2002). For example, although there may be substantial disagreements about philosophies and strategies among the advocates for comprehensive reforms (e.g., Success for All and Accelerated Schools), they are in a coalition of advocates for comprehensive reform. Also, with reading reform, there are long-standing disputes about direct instruction, but many reforms aligned with a literature-rich philosophy (e.g., Reading Recovery) benefit from the new advocacy for direct instruction, which has influenced new state and federal investments in reading reform. Thus, all three rationales influence state and federal reform efforts. Indeed, all three rationales have been used in states to promote early reading reform.

Once they are generally accepted, policy rationales can have a pervasive influence on education policy, but the underlying claims may move along unexamined by educators, evaluators, and policy makers. This article treats these rationales as reform claims and uses the analyses of surveys of early reading programs in Indiana's elementary schools to examine evidence related to these hypotheses. This explanatory analysis examines how

evaluation research on program outcomes can inform the evolution of policy rationales. We describe the three rationales, present the study approach used to collect and analyze results, discuss the statistical analyses, reconsider the hypotheses in relation to the empirical evidence, and summarize by considering the implications of the study for research and educational policy.

REFORM CLAIMS

Although it is widely recognized that different theories, or hypotheses, guide reform efforts, researchers seldom investigate competing rationales for reform. To achieve this goal, it is necessary to discern the assumptions made by the reform advocates, as well as to analyze data that permits an objective testing of these assumptions. This is a complicated task because most of the research on reading reforms holds implicit advocacy positions. Thus, not only is it difficult to conceptualize research in ways that make it possible to assess different assumptions about reform, but it can be even more difficult to find data that can be used to test these assumptions.

Instead of testing the rationales for reforms, most researchers conduct confirmatory research related to a specific reform model. By testing a single hypothesis rather than examining empirical evidence related to the specifics, research reports play the role of providing confirmatory evidence for individual reforms. However, the claims made as part of the rationale for funding often go untested. To examine the effects of rationales, we must start by examining the political arguments that influence the evolution of a funding program.

This study examines the effect of Indiana's early reading reform program on three school outcomes (special-education referral for Grades K-3, retention rates for Grades K-3, and pass rates on the state third grade reading achievement test). This article examines 3 years of surveys from schools funded in the Indiana Early Intervention Grant Program (EIGP), along with surveys from comparison schools. Below, we outline the three policy rationales and consider how they influenced the EIGP.

Professional Development

The professional development rationale essentially is that teachers need time for professional development that involves collaboration on strategies for improving educational outcomes. In the mid-1990s, the Indiana Department of Education (IDOE) initiated efforts to support professional development for teachers. Initially, they studied state policies on professional devel-

opment, focusing on release time for teachers in various states (Bull, Buechler, Didley, & Krehbiel, 1994). The IDOE has continued to advocate for subsidizing release time for teachers' professional development. As a second phase of their advocacy process, the IDOE commissioned a study of the research base for professional development and found that "successful" professional development is

- school based, that is, focused on particular problems of each school and selected by the teachers and principal to help address those problems;
- followed up in teachers' classrooms by such means as peer coaching or school-based research and evaluation teams;
- collaborative, so that groups of professionals at schools can work together to solve the school problems most critical to student learning;
- embedded in the daily lives of teachers, so that they undertake professional growth as a central responsibility of teaching; and
- focused on student learning, so that teams of teachers at the school attend to the actual effects of professional development on the performance of the students for whom they are responsible. (Bull & Buechler, 1996, p. 5)

These principles have been widely disseminated in Indiana as well as nationally. In Indiana, the IDOE strongly encourages schools to consider these principles whenever they submit proposals for categorical funds (St. John, Ward, & Laine, 1999). Therefore, the professional development rationale was embedded in the program and was integral to the funding process in EIGP. To receive a grant, a school either received a subsidy for professional development in Reading Recovery or needed to include a focus on professional development (consistent with the principles above) in their grant application.

Unfortunately, few conceptual frameworks have been developed for the analysis of the effects of professional development on student outcomes. These conceptual models argue that professional development has a direct influence on professional practices and influences student outcomes as a result of changes in educational practice outcomes (Guskey & Sparks, 1997; St. John et al., 1999). This conceptualization is consistent with the principles outlined above and thus had an influence on the conceptual model for this study. We included an explicit examination of professional development activities as part of our surveys.

Comprehensive Reform

A second reform rationale is that comprehensive and cohesive reform approaches are essential to improving schools' outcomes. In the debates over

reading intervention strategies, Barbara Taylor and her colleagues (2000) have argued this position. In recent years, the comprehensive reform argument has also had a substantial influence on the Comprehensive School Reform Demonstration and schoolwide Title I (Wong, 2003). This reform argument had a substantial influence on EIGP in Indiana. In particular, advocates of one program (Reading Recovery¹) influenced the creation of EIGP in Indiana. Reading Recovery was made the first priority for funding. Over time, about half of the funds were used for Reading Recovery and about half for other programs.

In addition to the EIGP funding, the schools in Indiana had access to other types of funding from other state and federal sources, so it was possible that comparison schools included in these surveys had similar reforms funded by alternative programs. Therefore, it was necessary to review the features of diverse reform models before we could develop a survey that could be used to analyze the effects of the different types of reform models that schools might use. The reforms that are represented in the current study are described briefly below.

Reading Recovery: A pullout, one-on-one reading intervention for the lowest achieving students in first grade (i.e., the lowest 20%). Reading Recovery (RR) is designed to bring those students up to grade level. The intervention helps children make the difficult transition from decoding to comprehension. The program has been widely studied by proponents and independent researchers (e.g., Center, Wheldall, Freeman, Outhred, & McNaught, 1995; Iverson & Tunmer, 1993; Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994; Rowe, 1997).

Success for All: A comprehensive school-restructuring process designed for schools with large populations at risk for learning failure. Success for All (SFA) balances a skills-oriented instructional approach with a heavy emphasis on collaboration and teamwork among educators. The program has a substantial research base (Dianda & Flaherty, 1995; Madden, Slavin, Karweit, Donlan, & Waskik, 1991; Madden, Slavin, Karweit, Limermon, & Donlan, 1989; Ross & Smith, 1994) mostly conducted by advocates for SFA.

Literacy Collaborative: A schoolwide restructuring model that focuses on classroom-based instruction, depending on Reading Recovery as a "safety net" for those students still not succeeding. Developed by the Reading Recovery program at Ohio State, the Literacy Collaborative (LC) provides support for students who are not receiving RR. LC involves the whole school—especially teachers and families—in a comprehensive and reflective approach to literacy instruction, which is appropriate for all children. This program is relatively new and lacks a confirmatory research base (for more information, see Ohio State University, Reading Recover Project, 1998).

Full-Day Kindergarten: Full-day kindergarten (FDK) is not funded by the state of Indiana as part of the state formula. Based on research-based models (e.g., Elicker & Mathur, 1997; Humphrey, 1988; Sheehan, Cryan, Wiechel, &

Bandy, 1991), EIGP funded FDK programs that provided enhanced approaches to literacy and math.

First Steps: A classroom-based language development model that serves as a teacher resource for closing the loop between diagnostic observation of child development and classroom instruction. First Steps provides teaching strategies, specific outcomes, and parent involvement ideas for each stage of child development. The research base is modest and mostly descriptive (Australian Council for Educational Research, 1993a, 1993b; Deschamp, 1995).

Even Start: An early intervention program that aims to help break the poverty cycle by improving educational opportunities for low-income families. Even Start is not intended to affect early reading directly but to create a developmentally appropriate home environment. A few federal studies have assessed the effects of the program, indicating it helps students make normal educational progress (e.g., Connor-Tadros, 1996; Gamse, Conger, Elson, & McCarthy, 1997).

Accelerated Schools Project: A schoolwide reform model, Accelerated Schools Project (ASP) is based on the notion that students in at-risk situations can learn at an accelerated pace by offering enriched curricula and instruction similar to that used for gifted education. ASP was designed as an inquiry-based professional development model with a clearly articulated philosophical base, which encourages active and reflective experimentation and evaluation. The program has a substantial research base (e.g., Finnan, St. John, Slovacek, & McCarthy, 1996), but the research that considers student outcomes is limited (e.g., Knight & Stallings, 1995; McCarthy & Still, 1993).

Four Blocks: Multilevel, multimethod instruction provided in a framework that provides an organized, systematic structure for providing early literacy instruction. The four "blocks" of the program are guided reading, self-selected reading, writing, and working with words. The model was proposed by Cunningham (1991) and has a modest research base (Cunningham, Hall, & Defee, 1991, 1998) conducted by the model advocates.

Direct Phonics Instruction

The third reform rationale is that direct instruction in letter/sound relationships (i.e., direct phonics instruction) can improve early reading acquisition. The most influential study on phonological awareness and the alphabetic principle is by Barbara Foorman and her colleagues (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998), a study that was independent of specific reform models. Recently, in response to critics, Foorman and her colleagues summarized their argument as follows:

However, we do maintain that there are some instruction principles that teachers and schools can use to enhance the reading achievement of at-risk children, and that it makes sense to demonstrate effective implementation of these instructional principles before investing in more complex solutions. (Foorman et al., 2000, p. 27)

Thus, although Foorman and her colleagues do not overtly argue against taking comprehensive approaches to early reading improvement, they do argue for a specific method as a first priority. This argument was based on research that found the following:

Controlling for differences in age, ethnicity, and verbal IQ, we found that children in the direct code (DC) approach improved in word reading at a faster rate and had higher word recognition skills in April than children receiving the implicit code (IC) approach (either research-based IC or district's standard IC). More importantly, children in *all* instructional groups with higher phonological processing scores in the beginning of the year demonstrated improvement across the year. (Foorman et al., 2000, p. 29)

Thus, this argument rests on research that shows that students who have instruction in direct coding learn to read words faster and that students who have skills in phonological processes learned to read faster. Some states have actually required phonics instruction in teacher education based on this line of research. The Snow report (Snow et al., 1998) concludes that schools should first emphasize direct-instructional approaches. However, it is not clear whether the particular reform models they advocated (e.g., Success for All) actually influenced schools to place more emphasis on direct instruction. Several states—including Texas, Washington, California, New York, and Wisconsin—have required direct phonics instruction (Allington & Woodside, 1999; Taylor et al., 2000). This approach to policy takes the claim made by Foorman et al. (1998) quite literally. That is, these researchers essentially claim that direct phonics instruction is a necessary first step, and legislators in these states have taken action requiring this instructional approach. The implicit political claim made in this approach is that requiring all schools to use an explicit approach to phonics instruction will increase the number of students who learn to read and make normal educational progress.

These arguments have had an influence on the evolution of the EIGP in Indiana in the sense that direct instruction is widely advocated. Many schools in Indiana used direct approaches in their early reading instruction. Furthermore, the survey instrument developed for this study allowed us to measure the extent to which schools were using direct instruction and related classroom practices. Thus, it was possible to test the direct-instruction hypothesis.

Examining Research Claims

The EIGP was implemented in fall 1997 and has been funded for the past 4 years. The program was funded by the state at \$4 million per year during the 3 years examined in this study. Three years of surveys (1997-1998, 1998-1999,

and 1999-2000) were conducted of schools funded by EIGP along with a random sample of schools that did not receive funding. However, funding was not random. Schools had to apply for funds. The surveys asked questions about actual classroom practices in both categories of schools and provided sufficient data to examine the effects of the three approaches to educational improvement on the following:

- two outcomes related to attainment/equity (rates of special-education referral and retention in Grades 1 through 3) and
- one outcome related to achievement (the percentage of third grade students passing the state reading test, the Indiana Statewide Testing for Educational Progress [ISTEP+], in the fall term after the intervention).

Specifically, we examined the effects of school characteristics, funding, type of intervention, professional development features, parent involvement features, and classroom practices (factor scores) on these outcomes. A factor that combined practices related to direct instruction was included in the regression models. These analyses provided a relatively comprehensive assessment of the effect of these reform models, which are examples of comprehensive reforms. Furthermore, because the survey instrument asked questions about professional development and direct instruction, it was possible to analyze information related to all three reform claims.

This analysis of the claims embedded in reform rationales is exploratory. The idea that the results of the evaluation provided evidence related to these sets of claims emerged after the formal evaluation reports were completed. In this study, we examine the influence of variables related to each set of claims (see Table 1). A set of dichotomous variables was used to examine the effects of having a specific type of reform on the student outcomes. Program features related to professional development were included in the questionnaire. These were coded dichotomous variables because survey respondents indicated only whether these practices were used during the school year of the survey. The questions about classroom practices asked about the frequency of use, using Likert-type scales. Classroom practices related to direct instruction converged as a factor. Factor scores for direct instruction and related practices were used in the analyses as an indicator of the extent to which practices related to direct instruction were used.

These measures represented proxies for the three sets of policy claims. To measure the effects of these variables on educational outcomes, we also needed a model that controlled logically and statistically for other forces that could have influenced these outcomes. Below, we describe our approach to that challenge.

TABLE 1
Proxy Measures Related to Policy Claims About Reform Strategies

<i>Claim/Variable</i>	<i>Variable Type</i>	<i>Relationship to the Claim</i>
Comprehensive reform claims (model type)	Dichotomous variables	These variables compare schools with specific types of interventions with other schools that did not have the reform.
Reading Recovery		Information of whether the school had the reform derived from school survey.
Four Blocks		
Success for All		
Literacy Collaborative		
Full-Day Kindergarten		
First Steps		
Even Start		
Accelerated Schools		
Professional development	Dichotomous variables	Survey response to questions about whether these services were available to teachers in Grades 1-3.
Certified trainer		
Certified specialist		
In-service workshops		
Networking		
Opportunity to collaborate		
Direct instruction and related methods.	Factor score	Survey responses using Likert-type scale on program features loaded on to this factor. One of multiple classroom factors. Factors indicated "perceived patterns of classroom practice."
Program features that loaded on factor:		
Basal readers		
Phonics instruction		
Reading drills		
Worksheets/workbooks		

RESEARCH APPROACH

Three years of survey data from schools in Indiana were collected as part of a comprehensive evaluation of the Early Intervention Program in Indiana. Below, we describe the survey we used, the logical model used in the study, and the instruction and related factors developed from the study of the entire 3-year database.

Survey and Response

The survey instrument, the Early Literacy Intervention Survey,² included questions about the school, types of interventions used, the features of the early reading program, the number of students referred and retained, and enrollment information that could be used to impute special-education referral and retention rates. In addition, we had access to a state-level database with information on test scores.

The survey assessed the frequency of use of nine organizational and structural features (ability grouping, basal readers, child-initiated learning centers, independent reading, one-on-one tutorials, pullout instruction, small groups, systematic evaluation, and trade books). It also assessed the frequency of use of 10 classroom instructional methods (Big Books, cooperative learning, creative writing and/or essays, drama, emergent spelling, paired reading, phonics, reading aloud, reading drills, and worksheets/workbooks). For these frequency-of-use questions, survey participants were asked to respond on a 5-point scale for the extent of use from 1 for *never* to 5 for *every day* for both the current year and the prior year by grade level (K, 1, 2, 3). The survey also asked whether five types of professional development processes (certified training, certified specialist, in-service workshops, networking, and opportunity for collaboration) and five features related to parent involvement (book distribution, family literacy, paired reading, parent conference, and parent volunteers) were used in kindergarten through Grade 3.

This study reports analyses for 3 years of surveys of funded and comparison schools. The overall response rate across the 3 years was 61%. Comparison schools were half as likely to be surveyed as funded schools. Therefore, comparison schools were weighted by 2 to adjust for the probability of being surveyed. Table 2 presents the response rates for each year.

Surveys were sent to principals for each school in the sample. In some instances, reading specialists in the schools may have completed the survey forms, but the principals were responsible for responding on the surveys. The response rates varied across the 3 years of the study (61% in 1999-2000, 43% in 1998-1999, and 57% in 1997-1998). We realize the evaluation program

TABLE 2
Number and Response Rate of Surveyed Schools

	<i>Funded</i>	<i>Comparison</i>	<i>Total</i>
1997-1998			
Number surveyed	262	351	613
Number responded	167	182	349
Rate (%)	64	52	57
1998-1999			
Number surveyed	289	359	648
Number responded	170	108	278
Rate (%)	59	30	43
1999-2000			
Number surveyed	186	373	459
Number responded	147	133	280
Rate (%)	79	35	61

was conceived of as a school-level study, so a survey was developed for each of the years included in this article. Representativeness of the survey could be limited by the fact that survey respondents were used in each school, as well as by subsequent variability in response rates. Therefore, the current study is appropriately characterized as exploratory. The response rates were higher for funded schools than for comparison schools.

Statistical Methods

The study used descriptive statistics, factor analysis, and multiple regression. The descriptive statistics describe the population characteristics.

A factor analysis was run for 19 variables related to instructional and classroom program features for the entire population. Specifically, the average Likert-type scale score for Grades 1 through 3 was imputed for the 19 program features for instruction and structural/organization features of the survey instrument. A conservative factor-loading minimum of .50 was used. Missing items were replaced with mean values.

Ordinary least squares (OLS) regression was used to examine the influence of predictor variables on the three outcomes. We present R^2 plus three levels of significance likelihood (.01, .05, and .10) for each predictor variable. Because .10 is only moderately significant, we make note of this moderate association in the text so the reader will not place undue emphasis on this statistical relationship.

Model Specifications

A multiple regression model was developed to assess the effects of school characteristics, intervention type, professional development and instruction, and related factors on three outcomes: special-education referral rates for Grades K-3, retention rates for Grades K-3, and passing rates of third grade ISTEP+ reading. The analyses for the first two outcomes (i.e., referral and retention rates) considered 3 years of surveys. The analysis for one outcome (i.e., test pass rates) considered 2 years of surveys. The blocks of variables, added sequentially, were as follows:

- *School Characteristics:* Whether a school was funded, the average ISTEP+ score, the percentage of students receiving free or reduced lunch, the percentage of minority students, and school locale (urban and rural schools were compared with students in other locale types).
- *Intervention Type:* Reading Recovery,³ Success for All, Literacy Collaborative, Full-Day Kindergarten, First Steps, Even Start, Accelerated Schools,⁴ and Four Blocks were coded as dichotomous variables. Some schools had more than one of these reform models.
- *Professional Development:* Whether reading teachers are required to be certified, whether certified specialists are brought in for training sessions, whether in-service workshops were used, whether teachers networked with teachers in other schools, and whether teachers collaborated within the school on reading instruction were included in professional development.
- *Parent Involvement:* Book distribution, family literacy, paired reading (parent to child), parent conferences, and parent volunteers were included as dichotomous variables.
- *Instruction and Related Factors:* We included the scores for the nine structural/organizational factors and the 10 classroom instructional methods.

The regression analyses stepped in blocks of variables that were related to each of these constructs (i.e., bullets above). This approach allowed us to see what additional understandings could be gained from each step. Specifically, we considered changes in the significance of independent variables across the sequence of steps, which revealed confounding relationships, providing more clues about effects. We also considered whether the addition of variables added to the explanatory power of the analysis, as indicated by change in the R^2 .

Instructional and Related Factors

Because of the large number of program features related to instruction and the organization of reading programs at the grade level, we decided to conduct a factor analysis of the instructional and structural/organizational fea-

tures. (The factor analysis is presented in St. John, Manset, Chung, Simmons, & Musoba, 2000).

- *The connected-text approaches* factor includes independent reading, cooperative learning, creative writing, emergent spelling, paired reading (student-to-student), and reading aloud. Schools that make use of these methods combine techniques that engage students in the learning process.
- *The direct/explicit approaches* factor combines basal readers, phonics instruction, reading drills, and worksheets/workbooks. Schools that use direct/explicit approaches emphasize systematic approaches to teaching the components of language and reading.
- *The child-centered/expressive approaches* factor includes child-initiated learning centers, Big Books, cooperative learning, and drama. These instructional approaches place an emphasis on the development of the whole child and peer engagement among children.
- *The small group/tutorial approaches* factor combines ability grouping, one-on-one tutoring, pullout instruction, and small groups. Schools that use these techniques place more emphasis on classifying children and accelerating the learning of some while addressing developmental needs of others.
- *The trade-books approaches* factor combines trade books and Big Books, but de-emphasizes basal readers. In this approach, schools use texts that are literature-based and engaging for students, rather than structured elements of reading programs that emphasize increasing levels of difficulty.

Of these factors, the variables included in the direct/explicit approaches factor are closely aligned with the direct approaches advocated by Foorman and colleagues (Foorman et al., 2000; Foorman et al., 1998). The other factors would seem more closely aligned with the more comprehensive approaches advocated by Taylor and other reform advocates. However, because these factors were developed from principals' responses to questions about classroom practices, we are cautious about our interpretation of these factors.

Limitations

This study has limitations that merit consideration by readers. First, our analyses consider school-related outcomes rather than individual outcomes. Whereas most reading research focuses on individual students, we felt it was important for the funding agency to understand whether their funding influenced student outcomes at the school level. This approach is consistent with the ways school outcomes are frequently reported to legislators and the public. Thus, this approach was appropriate for a policy study of this type.

Second, the survey asked respondents to answer questions about program features for all grade levels rather than asking each teacher to respond to a questionnaire. We considered this approach appropriate for this initial test of

the study methodology. In the future, we plan to extend the method to include a survey of teachers, which would mean we could examine both school-level and classroom-level outcomes.

Third, we assumed the schools that responded to the survey were representative of all of the schools in the funded and comparison groups. This assumption was necessary because of the statistical methods, but it seems a reasonable assumption for this study population. Although funded schools and comparison schools had different response rates, these differences seem reasonable, given that funded schools had more motivation to respond as a result of their involvement in the program.

Fourth, although many of the comparison schools in Indiana used direct-instruction methods, the funded interventions did not place a substantial emphasis on these practices, an issue considered below. In the summer of 2000, the IDOE provided statewide training, developed a Web page, and developed Web-based courses (for continuing education courses) that emphasized direct instruction. In the next study, it may be possible to further test this hypothesis.

Fifth, the factor analysis used the principals' responses to questions about classroom practice. For most of the variables in the study, a school-level measure was used, so the principals' responses were appropriate. However, for the classroom practices, principals could only express their opinions about the practices that were being used and the extent of use. Therefore, we are cautious in reaching conclusions about the classroom practices. The surveys represent opinions rather than observation of practices. Because our surveys were filled out by principals or teachers who were familiar with the reading programs, they were probably informed by observation. However, at the time of response, those filling out the questionnaires were giving their opinions, often using Likert-type scales. So, although these analyses provide information about perceptions of implementation, they are not reporting actual observations of classroom practice.

Sixth, as noted in Table 2, there was variation in response rates across years, as well as for treatment schools in relation to comparison schools within each year. We used sample weights as a partial adjustment for the variable in response rates within years for the comparison and treatment schools. In future studies, it would be appropriate to use STATA or other statistical methods that correct for the use of cross-year surveys. However, the use of multiple years of surveys enabled us to survey every elementary school in the state, which improves our ability to treat these analyses as representative of the state as a whole.

Finally, the measurement of effects in the analyses presented in this article represent the direction of association between certain types of projects—and

related program features—and the outcome they intend to influence. Although some of the associations examined in the models present in this article are statistically significant, these analyses do not provide estimates of the number of students affected by these interventions. The actual number of students affected by these interventions might be quite small. The intention of this project is not to measure effects but rather to explore measurable patterns of association as a means of discovering whether policy claims are reflected in the results of well-designed, quasiexperimental evaluations.

FINDINGS

Sample Characteristics

The characteristics of the Indiana schools included in this study are presented in Table 3. The three outcomes provide very different measures of school success. The average rates for special-education referral (5%) and grade-level retention (2%) were relatively low. In contrast, an average of 69% passed the ISTEP+ reading/language arts test.

The schools in this study represent a great diversity in school characteristics. The responding schools had an average of 26% of free or reduced lunch students and an average of 13% minority students. About one fifth were from city locales (18.5%) and about one-third rural (31.9%). The other half were located in suburban and town locales. In addition, about one third of the sample received 1 year of funding (32%), and 11.3% received 2 years of funding, with very few schools receiving funding all 3 years.

The sample population was reasonably reflective of the state population of schools, and the sample analyses of school characteristics are compared with state averages for all schools in Table 3. The percentage of students passing reading tests and raw scores for these exams were identical for the sample and the population. The percentage of students in the federal Free and Reduced Lunch Program, our measure of poverty, was higher for the sample than for the population. However, the percentage of minority students was lower for the sample than for the populations. The response rates for urban and rural schools were lower than for schools in other locales (towns and suburbs).

Three types of interventions represented the majority of comprehensive interventions: Reading Recovery (40.8%), Full-Day Kindergarten (12.3%), and Four Blocks (12.5%). These methods were encouraged by the IDOE through different initiatives during the period of the study. Local political forces influenced the number of schools that applied for state funds to

TABLE 3
Descriptive Statistics of the Sample

	<i>Mean (%)^a</i>	<i>State Average 1999 (%)</i>
Outcome variables		
Special education Grade 1-3	5	
Grade retention Grade 1-3	2	
% Passing ISTEP English/language arts scale score	69	69
School characteristics		
ISTEP reading raw score ^b	34.44	34.27
% Free or reduced lunch	26	20
% Minority	13	25
City ^c	18.5	
Rural	31.9	
EIGP funding status		
1-year funding	32.0	
2-year funding	11.3%	
3-year funding	1.3	
Intervention type ^d		
Reading Recovery	40.8	
Success for All	1.5	
Literacy Collaborative	3.2	
Full-Day Kindergarten	12.3	
First Steps	3.0	
Even Start	0.9	
Accelerated Schools	0.7	
Four Blocks	12.5	
Professional development		
Certified training	32.7	
Certified specialist grade	34.4	
In-service workshops	77.0	
Networking	66.1	
Opportunity for collaboration	73.5	
Parent involvement		
Book distribution	50.9	
Family literacy	30.5	
Paired reading (parent-to-child)	76.2	
Parent conferences	97.3	
Parent volunteers	64.2	

NOTE: $n = 823$; double weight was given to comparison schools. ISTEP = Indiana Statewide Testing for Educational Progress; EIGP = Early Intervention Grant Program.

a. Percentages only are reported for dichotomous variables. Averages and standard deviations are reported when percentages are used as continuous variables.

b. These figures are not percentages; they are mean scores.

c. Schools in town and suburban locales are the reference group.

d. Schools having no or other interventions were the reference group.

implement these methods.⁵ Much smaller percentages of the schools used Success for All, Literacy Collaborative, and other methods. However, this list of reforms does not correspond perfectly with receipt of funds. For example, not all Accelerated Schools received funding through the EIGP.

Most schools in the sample used in-service workshops (77%), networking (66.1%), and collaboration (73.5%). Certified training and certified specialists, professional development methods linked to Reading Recovery in Indiana, were used in about one third of the schools in the sample.

Parent involvement was widely used as a strategy in early reading programs. About half of the schools used book distribution. Most used parent conferences, parent volunteers, and paired (parent-to-child) reading. About one third had family literacy programs. Table 3 also compares the demographic characteristics of schools responding to the surveys with the characteristics of schools in the state.

Referral Rates

The analysis of referral rates, represented as a sequential set of regression models (Table 4), reveals that all but one of the variables that were significantly associated with special-education referral had a consistent effect across all of the models. Each set of variables is considered below.

First, two variables related to school characteristics had a significant association with special-education referrals across all of the models. Having high average ISTEP+ reading scores the prior year was negatively associated with referrals, indicating that schools with higher test scores had lower referral rates, controlling for other school characteristics. Furthermore, the percentage of students on free or reduced lunch was positively associated with special-education referrals, indicating a relationship between poverty and special learning needs. None of the other school characteristics was significantly associated with referral. School characteristics explain only 4.3% of the variance in referral rates.

Second, four of the intervention types had consistently significant associations with referral rates. Schools with Success for All, Literacy Collaborative, and First Steps were negatively associated with special-education referral rates. Controlling for other variables in the model, these interventions help more students stay in the educational mainstream. In contrast, Four Blocks was significant and positively associated with referral rates. Step 2 adds modestly to the explanation of variance, raising explained variance to 6.2%.

Third, none of the professional development variables were significant in any of the analyses of special-education referrals. This indicates that professional development for reading was not associated with referral rates.

TABLE 4
Standardized Coefficients of Predictors on Special-Education Referral Rate

<i>Variable</i>	<i>Beta</i>				
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>	<i>Model 5</i>
School characteristics					
1-year funding index	-0.024	-0.024	-0.027	-0.027	-0.032
2-year funding index	-0.024	-0.035	-0.036	-0.032	-0.042
3-year funding index	-0.002	-0.010	-0.011	-0.008	-0.014
ISTEP reading total raw score	-0.123***	-0.129***	-0.127***	-0.133***	-0.125***
% getting free lunch	0.189***	0.181***	0.183***	0.191***	0.186***
% minority	-0.065	-0.044	-0.047	-0.052	-0.043
City	-0.024	-0.029	-0.030	-0.031	-0.031
Rural	0.019	0.033	0.032	0.034	0.027
Intervention type					
Reading Recovery		0.028	0.018	0.019	0.017
Success for All		-0.064*	-0.065*	-0.068*	-0.075**
Literacy Collaborative		-0.081**	-0.078**	-0.078**	-0.092**
Full-Day Kindergarten		0.050	0.046	0.045	0.043
First Steps		-0.081**	-0.078**	-0.078**	-0.087**
Even Start		-0.022	-0.019	-0.024	-0.029
Accelerated Schools		0.009	0.006	0.004	0.007
Four Blocks		0.080**	0.084**	0.085**	0.072*
Professional development					
Certified training			0.046	0.050	0.041
Certified specialist grade			0.035	0.037	0.023
In-service workshops			-0.005	-0.003	-0.012

(continued)

TABLE 4 (continued)

<i>Variable</i>	<i>Beta</i>				
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>	<i>Model 5</i>
Networking			-0.033	-0.029	-0.028
Opportunity for collaboration			-0.009	-0.006	-0.019
Parent involvement					
Book distribution				-0.027	-0.032
Family literacy				-0.052	-0.065*
Paired reading				-0.008	-0.015
Parent conferences				0.009	0.006
Parent volunteers				0.029	0.022
Classroom practices					
Connected text					0.065*
Direct/explicit					-0.051
Child-centered/expressive					0.023
Ability group/pullout					0.076**
Trade books					0.036
Adjusted R^2	0.043	0.062	0.060	0.058	0.064

NOTE: $n = 823$; ISTEP = Indiana Statewide Testing for Educational Progress.* $p \leq .1$. ** $p \leq .05$. *** $p \leq .01$.

Furthermore, the R^2 actually declined in the third step, further indicating that professional development variables do not add to the explained variance in referral.

Fourth, one of the parent involvement variables, family literacy, was modestly significant (.10 alpha) in the final step but was not statistically significant when the instructional factors were taken into account. Controlling for all variables in the model, family literacy was associated with lower referral rates. This indicates that family literacy programs help families and schools to work together in a complementary way to keep more children in school. Again, the R^2 drops, indicating parent involvement is not associated with referrals.

Fifth, two of the factors related to classroom practices were significant. Connected-text approaches and ability grouping/pullout approaches were significant and positively associated with special-education referral. Schools that emphasized these practices had higher referral rates, controlling for other variables in the model. The R^2 improves modestly in the last step. This indicates that principals' perceptions of practice, as measured by these factors, have an association with the referral outcome.

Readers are cautioned not to leap to any conclusions about program efficacy from these analyses. The fact that differences in the rate of referral were associated with different reforms and the various features that are related to these reform models merits note. However, the amount of variation actually explained by the model is quite modest. Therefore, changing classroom practices in reading instruction could possibly reduce special-education referral in a few instances, but it is not a remedy for the forces that influence the need for special education. Rather, these analyses were designed to test claims often made by reforms—we find associations, controlling for other forces, but they explain a very small portion of the variance.

Retention Rates

The analyses of retention rates reveal a complex pattern of confounding relationships between predictor variables, as well as a number of direct relationships between predictor variables and retention rates (Table 5). Each set of variables is examined below.

First, five variables in the first model were significant in at least one version of the model. Two of the variables—the percentages of students getting free or reduced lunch and of minority students—were significant and positively associated with retention across all of the models. Schools with more low-income and minority students had more retention. School characteristics

explain 10.9% of the variance in grade-level retention. This is far better than the previous model but is still only a small portion of the variance.

Two of the funding variables—receiving 1 year of funding and receiving 2 years of funding—were significant and negatively associated with retention in the first two steps of the analyses. Thus, controlling for other school characteristics, receiving 1 or 2 years of funding through EIGP was associated with keeping more students on grade level. Both variables were highly significant (.05 alpha) in the first step, modestly significant in the second step (.10 alpha), and not significant in any of the subsequent steps. The drop in the extent of significance across both steps indicates that the influence of funding is related to the variables in the each of these steps. The significance of the two funding variables was mitigated when the types of interventions were added in the second model. This is a logical finding given that the funding supported the programs that many schools implemented. Another portion of the significance of the funding variables was related to the professional development opportunities in the school. Specifically, when having an opportunity to collaborate was considered, the funding variables were no longer significant. This suggests that receiving EIGP funds enabled teachers to collaborate, another desirable outcome of funding, at least according to the professional development hypothesis.

One variable, being in a rural locale, was modestly significant (.10 alpha) after the program types were added to the model. This suggests a confounding relationship between intervention types and the rural locale. Further analysis of the types of programs implemented in rural locales would be needed to discern why there was such a confounding relationship.

Second, two variables related to intervention types were significant in at least one step of the model. The Literacy Collaborative was modestly significant (.10 alpha) and negatively associated with retention in the first two versions of the analysis. There is a confounding relationship between the Literacy Collaborative and variables related to parent involvement. Because the Literacy Collaborative encourages family involvement (e.g., book distribution) in reading, this was not an unexpected finding. The intervention type variable added only modestly to the explanatory power of the model, raising the R^2 to .111.

Reading Recovery was also modestly significant (.10 alpha) after professional development variables were added to the model (Steps 3, 4, and 5). Because Reading Recovery provides certified training and because schools with this program had certified specialists in the first grade, we expect the confounding relationship was attributable to these variables. Certified training and certified specialists were positive but not significant when they

TABLE 5
Standardized Coefficients of Predictors on Retention Rate

<i>Variable</i>	<i>Beta</i>				
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>	<i>Model 5</i>
School characteristics					
1-year funding index	-0.103***	-0.069*	-0.057	-0.051	-0.034
2-year funding index	-0.103***	-0.067*	-0.057	-0.053	-0.034
3-year funding index	-0.048	-0.024	-0.015	-0.013	-0.005
ISTEP reading total raw score	0.036	0.017	0.015	0.028	-0.002
% getting free lunch	0.204***	0.203***	0.206***	0.188***	0.173***
% minority	0.221***	0.227***	0.236***	0.239***	0.222***
City	-0.046	-0.044	-0.050	-0.050	-0.045
Rural	0.058	0.062*	0.069*	0.064*	0.067*
Intervention type					
Reading Recovery		-0.055	-0.067*	-0.078*	-0.069*
Success for All		-0.045	-0.046	-0.052	-0.053
Literacy Collaborative		-0.064*	-0.064*	-0.056	-0.049
Full-Day Kindergarten		-0.018	-0.021	-0.014	-0.024
First Steps		-0.010	-0.008	0.000	-0.004
Even Start		-0.008	-0.005	-0.006	-0.006
Accelerated Schools		0.019	0.020	0.032	0.032
Four Blocks		-0.029	-0.023	-0.021	0.011
Professional development					
Certified training			0.010	0.006	0.007
Certified specialist grade			0.006	0.006	0.009
In-service workshops			0.054	0.042	0.049

(continued)

TABLE 5 (continued)

<i>Variable</i>	<i>Beta</i>				
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>	<i>Model 5</i>
Networking			-0.005	-0.004	-0.014
Opportunity for collaboration			-0.113***	-0.111***	-0.108***
Parent involvement					
Book distribution				-0.014	-0.005
Family literacy				0.028	0.027
Paired reading				-0.094***	-0.085**
Parent conferences				0.089**	0.091**
Parent volunteers				-0.040	-0.053
Classroom practices					
Connected text					-0.082**
Direct/explicit					0.079**
Child-centered/expressive					0.111***
Ability group/pullout					0.001
Trade books					-0.062*
Adjusted R^2	0.109	0.111	0.118	0.126	0.145

NOTE: $n = 823$; ISTEP = Indiana Statewide Testing for Educational Progress.* $p \leq .1$. ** $p \leq .05$. *** $p \leq .01$.

entered the model. The positive association in these variables offset the negative association of Reading Recovery in the last three steps.

Third, only one professional development variable was significant across all of the models. Having the opportunity to collaborate was negatively associated with retention, indicating that when teachers have the opportunity to collaborate, it is easier for them to communicate about the types of educational programs that will enable more children to achieve on grade level. Professional development variables modestly improved the R^2 , raising the R^2 to .118.

Fourth, two variables related to parent involvement were significant. Paired reading between parents and children was associated with lower retention rates, whereas parent conferences were associated with higher retention rates. Paired reading is directly related to literacy instruction and is an integral feature of Literacy Collaborative, which could explain the association between this program type and professional development. The parent involvement variable also added to R^2 , raising it to .126. This indicates that parent involvement contributes to student educational progress.

Finally, four of the factors were significant. Connected-text approaches and trade books were significant and negatively associated with retention rates, indicating these patterns of practice help more children perform on grade level. Direct/explicit approaches and child-centered/expressive approaches were significant and positively associated with retention rates, indicating these methods were associated with having fewer students perform on grade level. The R^2 also increased (to .145), indicating that there was a relationship between principals' perceptions of classroom practice and keeping students on grade level.

Although these analyses indicate measures of association between specific types of interventions and retention rates, these analyses are not intended as estimates of the number of students passed as a result of the intervention. We expect these numbers are quite small, given the small amount of variance explained by this model. However, these analyses do indicate patterns of association that should be of general interest.

Test Pass Rates

The analyses of the influence of predictor variables on pass rates for standardized tests (Table 6) also indicated a complex pattern of relationships between predictor variables. Three of the school characteristic variables had a consistent, significant relationship with the outcome across each version of the model. Schools with higher average ISTEP+ scores during the base year had higher pass rates the next year. In contrast, the percentages of minority

and low-income students were negatively associated with pass rates across all models. School characteristics explained more than half of the variances in test pass rates.

In this analysis, having 1 year of funding was significant and negatively associated with pass rates in the first two models. This variable was only modestly associated with pass rates in the third model (.10 alpha) and was not significantly associated with the outcome in the last two models. Thus, funding had a confounding relationship with both professional development (and especially the opportunity to collaborate) and parent involvement. This reveals the opposite pattern as the prior analysis. Whereas the combination of funding and the opportunity to collaborate seemed to help keep more children in the classroom, the similar combination of variables was associated with lower pass rates.

Three of the intervention types had significant associations with test pass rates. Success for All had a strong negative association with pass rates across all of the models. In contrast, both Literacy Collaborative and Accelerated Schools Project had modest (.10 alpha) positive associations with higher pass rates after parent involvement was considered. Literacy Collaborative and Accelerated Schools Project place a substantial emphasis on parent involvement. However, adding the type of intervention to the model modestly improved prediction (R^2 increased from .556 to .561).

Opportunity to collaborate was significant and negatively associated with pass rates. These findings reveal a confounding relationship between funding, professional development, and pass rates. However, adding professional development to the model added very little to the R^2 .

Two variables related to parent involvement were significant. Book distribution was positively associated with test scores. Schools that send books home with the children had more students pass reading tests. However, paired reading (parent/child) had a slight negative association with the outcomes in the last step. Both variables had only a modest association with the outcome (.1 alpha). Also, the block of variables added only modestly to the R^2 .

None of the factors related to classroom practices was significant. Prior analyses that had considered the type of program and type of program funding separately revealed a positive relationship between direct/explicit approaches and pass rates (St. John et al., 2000; St. John et al., 2003). The R^2 changed very little in the last step. The present model provides a better way to control for the effects of funding and program type. Therefore, controlling for funding and program type, we conclude that none of the patterns of classroom practice were associated with pass rates independent of the effects of the types of programs implemented and the receipt of funding.

TABLE 6
Standardized Coefficients of Predictors on Passing Rate in ISTEP+ English/Language Arts Scale Score

Variable	Beta				
	Model 1	Model 2	Model 3	Model 4	Model 5
School characteristics					
1-year funding index	-0.091***	-0.072**	-0.064*	-0.049	-0.045
2-year funding index	-0.015	-0.020	-0.017	-0.014	-0.011
3-year funding index					
ISTEP reading total raw score	0.523***	0.515***	0.521***	0.528***	0.522***
% getting free lunch	-0.198***	-0.203***	-0.198***	-0.208***	-0.228***
% minority	-0.148***	-0.158***	-0.154***	-0.158***	-0.146***
City	0.009	0.007	0.007	0.010	0.010
Rural	-0.021	-0.028	-0.020	-0.021	-0.018
Intervention type					
Reading Recovery		-0.034	-0.040	-0.046	-0.042
Success for All		-0.087***	-0.086***	-0.089***	-0.089***
Literacy Collaborative		0.046	0.047	0.054*	0.055*
Full-Day Kindergarten		0.021	0.016	0.016	0.019
First Steps		0.022	0.019	0.015	0.010
Even Start		0.024	0.028	0.033	0.033
Accelerated Schools		0.043	0.043	0.052*	0.051*
Four Blocks		0.010	0.015	0.024	0.027
Professional development					
Certified training			0.004	-0.002	0.002
Certified specialist grade			0.009	0.011	0.012
In-service workshops			0.042	0.040	0.040

(continued)

TABLE 6 (continued)

<i>Variable</i>	<i>Beta</i>				
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>	<i>Model 5</i>
Networking			-0.022	-0.021	-0.028
Opportunity for collaboration			-0.076**	-0.079**	-0.073**
Parent involvement					
Book distribution				0.060*	0.058*
Family literacy				-0.031	-0.031
Paired reading				-0.048	-0.053*
Parent conferences				0.006	0.005
Parent volunteers				-0.019	-0.019
Classroom practices					
Connected text					0.034
Direct/explicit					0.051
Child-centered/expressive					-0.017
Ability group/pullout					-0.023
Trade books					-0.021
Adjusted R^2	0.556	0.561	0.564	0.565	0.566

NOTE: $n = 535$; ISTEP = Indiana Statewide Testing for Educational Progress.* $p \leq .1$. ** $p \leq .05$. *** $p \leq .01$.

Although the amount of variance in test scores is higher than for the other models, demographic background and historical achievement explain a substantial amount of variance. The point of this analysis is to illuminate patterns—in addition to testing claims about the effect of reforms—so readers are reminded that only small amounts of variation in test pass rates were attributable to any of the programs studied.

Looking Across Models

When we look across the models, we find that very little variance in special-education referral was explained. The analysis of retention was slightly better, but the R^2 was modest. Finally, the analysis of test pass rates explained more variance, but demographic variables provided the most powerful predictors of this measure of student achievement. These interventions had an apparent influence, but it was modest at best.

EVIDENCE RELATED TO REFORM CLAIMS

Three rationales influenced the development and evolution of the EIGP in Indiana. By considering the results above in relation to the three reform claims, it is possible to untangle further how government-sponsored interventions influence learning outcomes. However, we need to be cautious about these results given that the features of the interventions did not explain much variance in student outcomes.

Professional Development

The professional development claim was that teachers need time for professional development that includes collaboration strategies for improving educational outcomes. These analyses reveal a pattern of relationship between funding and the opportunity to collaborate. In two of the three analyses, the effects of receiving funding disappeared when the effects of the opportunity to collaborate were also considered. There are three distinct aspects of this pattern of findings that merit attention.

First, this pattern of findings seems consonant with the model of professional development (Bull & Buechler, 1996) that the state of Indiana has encouraged schools to emphasize in all categorical grant applications (St. John et al., 1999). This finding indicates that the state's encouragement of professional development had an influence on the strategies schools used to involve teachers in their reform efforts.

Second, the opportunity to collaborate was associated with less retention and higher test pass rates. These approaches help keep children in the mainstream and improve passing tests. Apparently, the collaboration can help schools transcend the tradeoff between retention and test-pass rates. When teachers have time to collaborate, they exhibit an increased capacity to keep more students achieving at grade level. However, there is an implied tradeoff between two sorts of outcomes—keeping children on grade level and having higher pass rates.

Third, 2 years of funding with a reading reform was also associated with both outcomes. Having 1 year of funding and having 2 years of funding were associated with lower retention rates, whereas only having 1 year of funding was associated with having lower test pass rates. Two years of funding provided more time for teachers to work through the problems associated with creating learning environments that help more children to achieve on grade level. If schools had 2 years of funding before children entered the third grade, children benefited from the programs. Having supplemental programs in both first and second grades apparently made a difference for students. Because children took the ISTEP+ test at the start of third grade, the 2 years of funding would have been sufficient to affect the reading programs prior to taking these achievement tests.

Not only do these three interpretations provide modest support for the professional development hypothesis, but they also suggest the possibility of a deeper pattern. One line of argument about school reform is that restructuring processes that involve teachers and parents can create caring learning communities that engage more children in active learning (St. John, Griffith, & Allen-Haynes, 1997). Not only do the findings above indicate a relationship between funding and opportunity to collaborate among teachers, but parent involvement also seems to be an important relationship with respect to test passing rates (see Table 6). “People” variables have the greatest effect, indicating more attention should be given to creating caring communities that support the development of children in schools. This pattern of relationships among variables merits further consideration in future studies.

Comprehensive Reform

A second reform claim was that comprehensive and cohesive reform approaches are essential to improving schools’ outcomes. This notion is not only a long-standing assumption of reforms such as Reading Recovery (Clay, 1993) but also a counterargument to newer claims about direct phonics instruction. The findings of this study reinforce the notion that

comprehensive and cohesive approaches to reform can influence educational outcomes, but it also reveals that the effects of reform models are far from monolithic.

The effects of the distinctive reform models are summarized in Table 7, along with interpretive comments related to these reforms. These analyses not only controlled for the effects of funding and school characteristics but also considered confounding relationships with professional development, parent involvement, and patterns of classroom practice. Different reform models had different effects, reinforcing the notion that not only do different comprehensive reform models have different features but they also have different effects. However, reform models only had modest effects on student outcomes, controlling for school characteristics and funding.

It is abundantly apparent that the reforms had different effects. One of the reform models, Literacy Collaborative, had the desired association with all three outcomes, indicating a comprehensive effect of student outcomes. Another model, Success for All, had a desired association with one outcome (special-education referral) and an undesired association with another (achievement tests), suggesting a tradeoff implied in selecting this model. This finding does not necessarily contradict other research on the model (e.g., Slavin & Fashola, 1998), because it takes more than 3 years to implement the model. A third model, Four Blocks, had only one undesired association (related to higher special-education referral). The other models noted in the table (Reading Recovery, Full-Day Kindergarten, and Accelerated Schools) had one desired outcome.

There is some evidence that the effects seemed directly related to the design features in these models. For example, the finding that Reading Recovery was associated with higher retention rates seems directly related to the program's focus on raising the reading level for first grade children who are having trouble learning to read. Even the lower-than-expected findings associated with Four Blocks seem related to the lack of emphasis on professional development and parent involvement in this model. Professional development was integral to the state reform strategy, and it had a direct effect on some outcomes, but Four Blocks per se was not positively associated with these outcomes. Clearly, schools should consider the design features of possible reform models when they are considering a reform strategy. However, as the measures of classroom practices were principals' perceptions rather than teachers' responses, we remain cautious about reaching any firm conclusions regarding classroom practices.

TABLE 7
Summary Analyses of the Effects of Reform Models: Significant Associations From Regression Analyses,
With Possible Explanations (Considering Confounding Relationships and Design Features)

<i>Program</i>	<i>Effect</i>	<i>Possible Explanations</i>
Reading Recovery	Associated with lower grade-level retention rates.	<ol style="list-style-type: none"> 1. Finding related to the effect of professional development (e.g., certified training and specialist). 2. Finding consonant with the design intent, which is to enable more children to achieve on grade level.
Success for All	<p>Associated with lower rates of special-education referral.</p> <p>Associated with lower pass rates on standardized tests for reading and language arts.</p>	<p>This finding is consonant with the program design, which encourages mainstreaming more children.</p> <ol style="list-style-type: none"> 1. Could be attributable to the mainstreaming of more children. 2. Long-term implementation may be needed to achievement improvement on standardized tests.
Literacy Collaborative	<p>Associated with lower special-education referral rates.</p> <p>Associated with lower grade-level retention rates.</p> <p>Associated with higher pass rates on standardized tests for reading and language arts.</p>	<p>Consonant with the program design, which intends to involve parents and children in paired reading.</p> <ol style="list-style-type: none"> 1. Finding related to the direct effects of professional development. 2. Provides a cohesive design that is linked to the professional development processes provided through Reading Recovery. <ol style="list-style-type: none"> 1. Finding related to the direct effects of parent involvement (i.e., book distribution). 2. Indicates a cohesive design, which builds on the strengths of schools and families.
First Steps	Associated with lower special-education referral.	Consonant with the design concept, which emphasizes systematically tracking the progress of children and intervening to promote learning outcomes.

Full-Day Kindergarten	Associated with lower special-education referral.	<ol style="list-style-type: none"> 1. Finding related to parent involvement (e.g., paired reading [parent-child]). 2. Consonant with the more comprehensive approach to Full-Day Kindergarten emphasized in the funding program. <ol style="list-style-type: none"> 1. Finding related to parent involvement. 2. Consonant with program design, which emphasizes involvement in school governance and in educational practices. 3. Consonant with program intent of accelerating the learning of children in situations that put them at risk.
Accelerated Schools Project	Associated with higher pass rates on standardized tests for reading and language arts.	<ol style="list-style-type: none"> 1. Model design lacks a consistent emphasis on parent involvement and ongoing professional development. 2. Other explanations for this problematic finding are not readily evident.
Four Blocks	Associated with higher special-education referral rates.	

Direct Instruction

The third reform claim was that direct instruction in letter/sound relationships (i.e., direct phonics instruction) can improve early reading acquisition. The findings presented here do not support the direct-instruction hypothesis. The effects of classroom practices were modest. Earlier analyses—earlier studies that did not control for the years of funding—found that direct/explicit approaches were associated with higher grade-level retention rates and with higher passing rates on standardized tests (Manset, St. John, Hu, & Gordon, 2002; St. John et al., 2000). These earlier studies did not examine the effects of funding and the types of reforms adopted but rather used treatment variables that combined funding and reform type. Thus, the current study has better logical and statistical controls for the effects of funding than the earlier studies, which probably explains the difference in findings (given that the funding variable has the reverse effect noted in these earlier studies). Thus, we conclude that the earlier finding—that direct/explicit approaches were associated with higher pass rates on standardized tests and lower retention—was an artifact attributable to the 1st year effects of funding (i.e., the 1st year of new practices). However, regardless of which set of findings one accepts, there is not strong support for the direct-instruction hypothesis.

The partial correlations between the program types of the factors provide more explanations for these findings (see Table 8). These analyses reveal that most of the comprehensive reform models we examined were correlated with factors other than direct/explicit approaches. These models emphasize literature-rich instruction and de-emphasize direct phonics instruction, worksheets, and basal readers. None of the reform models was significant and positively correlated with the direct/explicit approaches factor. In fact, schools with four of the reform models—Reading Recovery, Success for All, Literacy Collaborative, and Four Blocks—were actually negatively associated with the direct/explicit approaches factor.

This supports the notion that a balanced approach to reading instruction may be needed. Apparently, there really is a difference between the reforms, such as Success for All, that were endorsed by the National Research Council (Snow et al., 1998) and direct approaches to reading intervention that are now so widely advocated. Success for All had a negative correlation with the direct/explicit approaches factor, which includes more emphasis on phonics and related practices. Similarly, several of the other reform models had similar patterns of practice as Success for All. Indeed, schools with Success for All—along with schools with Reading Recovery, Literacy Collaborative, and other locally developed programs—were associated with connected-text approaches and ability-group/pullout approaches (Table 6). Thus, a

TABLE 8
Relationship Between Reading Reforms and Frequency of Patterns of Practice Reported by Schools

<i>Program</i>	<i>Text-Connected Approaches</i>	<i>Direct/Explicit Approaches</i>	<i>Child-Centered/ Expressive Approaches</i>	<i>Ability-Group/ Pullout Approaches</i>	<i>Trade Book Approaches</i>
Reading Recovery	+	-		+	
Success For All	+	-		+	-
Literacy Collaborative	+	-	+	+	
First Steps					
Accelerated Schools					
Four Blocks	+	-		-	+
Early Success					
Other locally developed programs	+		+	+	
No program	-			-	

NOTE: + = Significant and positive Pearson correlation; - = Significant and negative Pearson correlation.

comprehensive approach may provide a literature-rich environment and supplemental support for children who are having trouble learning to read. In the summer of 2000, the IDOE sponsored statewide training on a “phonics toolkit” and development of Web-based continuing education in methods for promoting phonemic awareness as a part of classroom instruction. Perhaps these workshops will introduce new methods in direct instruction that will have the desired influence on student outcomes.

However, our findings about instructional practices are appropriately viewed as exploratory, given the limitations of the surveys. Therefore, it may be too early to assess adequately whether the direct-instruction rationale holds up to empirical evidence. In Indiana, at least, research to date does not support the direct-instruction hypothesis. Therefore, we think further analyses are needed to test more fully whether these reform claims hold up when reforms are implemented that emphasize phonics and other direct approaches to early reading instruction.

CONCLUSIONS AND IMPLICATIONS

The primary conclusion that can be drawn from this study is that reform advocates often oversimplify the complexity of reforming reading programs. While the idea that a research base can inform policy is compelling, adhering to a policy claim based on research that was conducted by reform advocates may be shortsighted. Policy makers and educators should be more skeptical about the claims made by researcher/reformers. Although there is some compelling evidence to support the core aspects of the claims examined here, each set of claims was more complex than portrayed by the reform advocates. And the direct effects of reform models do not explain much variance in student outcomes. Thus, although the new wave of research-based reading reforms holds promise, it is far from a panacea for educators.

This study indicates it is important to consider the role of the various rationales (or policy arguments) that influence the evolution of a grant program when evaluating the effects of the program. In particular, we found that the state’s emphasis on embedding professional development in all grant proposals explained some of the direct effects of funding, after the types of interventions were controlled. This finding reveals that placing an explicit emphasis on professional development in the funding process can have a positive effect on students if the intervention is sustained for at least 2 years. Although specific comprehensive reform models also had effects, different models had different effects. This means that schools should choose program types that meet their needs—that is, address deficiencies in their existing programs.

Thus, the general rationale that comprehensive reforms improve student outcomes seems to overgeneralize a complex set of relationships between reforms and student outcomes. Furthermore, there is reason to raise questions about the practical meaning of the direct-instruction rationale. Specifically, it is difficult to discern how these practices actually influence student outcomes in schools, given that many schools already use these practices.

These findings have three important implications. First, it is crucial that policy makers include comprehensive evaluation studies when they undertake major new reforms. Too often, state education reforms do not include evaluations; nor do evaluations routinely test the assumptions made by reform advocates. Policy makers need to ponder the possibility of the null hypothesis. Indeed, the results of this study suggest that reform strategies that allow educators to select intervention designs that meet their educational needs may have a greater effect than mandating single reform models. For example, the national implementation of the Comprehensive School Reform Demonstration has enabled many schools to choose among possible reform models. This approach, which is similar to the approach used by the state program examined in this report, seems a more reasonable approach than advocating a specific reform method in the legislation for education reform.

Second, it is crucial that evaluators think critically about the claims made by various educational reformers. These claims should be evaluated and openly discussed. The rationales used by reformers are generally based on research that tests a single claim. Because the funding programs respond to a range of policy arguments in their design and evolution, it is important that researchers examine the effect of the specific activities that are required by a funding agency to secure funds. It also is necessary to control for the contexts in which reforms are implemented, as well as for classroom practices in comparison schools. In this study, we were careful to discern the features of various reform models and to ask educators about whether those features were frequently used in their schools. This enabled us to assess the effect of patterns of classroom practices on educational outcomes and to control for these patterns when we assessed the effect of reform models. When we took this step, it became apparent that different reform models had different effects. Furthermore, we found that the designs themselves helped to explain the effect of the reform models.

Third, it is important that educators take the opportunity to think critically about the features of various reform models and the ways these features link to educational outcomes and the ways the reform models would change their education practices. To do this, they need to be able to assess their current practices, to consider the features of current reading programs, and to compare these practices with the features of various reforms. Some work has

been done in producing guides that can help educators assess current practices (e.g., North Central Regional Education Laboratory, 2000; St. John & Bardzell, 1999), but more work is needed. Given that the new wave of research-based reading reforms is creating opportunities for educators to make informed choices about improvement strategies, a better effort should be made to disseminate useful information into the hands of educators who are making choices about reform strategies.

NOTES

1. Reading Recovery is further described below.
2. Copies of the survey can be obtained on request from the Indiana Education Policy Center. A simplified version of the instrument that can be used to survey teachers is available online (St. John & Bardzell, 1999).
3. This variable coding included schools with Reading Recovery regardless of whether they were funded through the Early Intervention Grant Program (EIGP).
4. The EIGP did not fund any Accelerated Schools, but this intervention type was discussed in documents disseminated through the program (St. John & Bardzell, 1999), and there were a few Accelerated Schools in the state.
5. These approaches were formally encouraged by the Indiana Department of Education.

REFERENCES

- Allington, R. L., & Woodside, H. (1999). The politics of literacy teaching: How "research" shaped policy. *Educational Researcher*, 28(8), 4-12.
- Australian Council for Education Research. (1993a). *The impact of First Steps on the reading and writing ability of western Australian year 5 school students* (An Interim Report to the Curriculum Development Branch Western Australia Ministry of Education). Perth, Australia: Western Australia Department of Education.
- Australian Council for Education Research. (1993b). *The impact of First Steps on schools and teachers* (An Interim Report to the Curriculum Development Branch Western Australia Ministry of Education). Perth, Australia: Western Australia Department of Education.
- Bull, B., & Buechler, M. (1996). *Learning together: Professional development for better schools*. Bloomington: Indiana Education Policy Center.
- Bull, B., Buechler, M., Didley, S., & Krehbiel, L. (1994). *Professional development and teacher time: Principles, guidelines, and policy options for Indiana*. Bloomington: Indiana Education Policy Center.
- Center, Y., Wheldall, K., Freeman, L., Outhred, L., & McNaught, N. (1995). An evaluation of Reading Recovery. *Reading Research Quarterly*, 30, 240-263.
- Chall, J. S. (1967). *Learning to read: The great debate*. New York: McGraw Hill.
- Clay, M. M. (1993). *Reading Recovery: A guidebook for teachers in training*. Portsmouth, NH: Heinemann.

- Connor-Tadros, L. (1996). *Effects of Even Start on family literacy: Local and national comparisons*. Baltimore: Johns Hopkins University.
- Cunningham, P. (1991). Research directions: Multi-method, multilevel, literacy instruction in first grade. *Language Arts*, 68, 578-584.
- Cunningham, P. M., Hall, D. R., & Defee, M. (1991). Non-ability grouped, multilevel instruction: A year in a first grade classroom. *Reading Teacher*, 48, 566-571.
- Cunningham, P. M., Hall, D. R., & Defee, M. (1998). Non-ability grouped, multilevel instruction: Eight years later. *Reading Teacher*, 51(8), 652-664.
- Deschamp, P. (1995). Case studies of the implementation of the First Steps project in twelve schools. Perth, Australia: Western Australia, Education Department. (ERIC Document Reproduction Service No. ED 419425).
- Dianda, M., & Flaherty, J. (1995). *Annual report on promising practices and program adaptations successes*. Los Alamitos, CA: Southwest Regional Educational Laboratory.
- Education Commission of the States. (1997). *Investing in teacher professional development: A look at 16 school districts*. Denver, CO: Putnam.
- Elicker, J., & Mathur, S. (1997). What do they do all day? A comprehensive evaluation of a full-day kindergarten. *Early Childhood Research Quarterly*, 12, 459-480.
- Finnan, C., St. John, E. P., Slovacek, S. P., & McCarthy, J. (Eds.). (1996). *Accelerated Schools in action: Lessons from the field*. Thousand Oaks, CA: Corwin Press.
- Foorman, B. R., Fletcher, J. M., Francis, D. J., & Schatschneider, C. (2000). Response: Misrepresentation of research by other researchers. *Educational Researcher*, 29(6), 27-37.
- Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, 90(1), 37-55.
- Gamse, B. C., Conger, D., Elson, D., & McCarthy, M. (1997). Follow-up study of families in the Even Start in-depth study: Final report. Cambridge, MA: Abt Associates, Inc. (ERIC Document Reproduction Service No. 413099).
- Gordon, W. E. (1999). *Education and justice: View from the back of the bus*. Albany: State University of New York Press.
- Guskey, T. R., & Sparks, D. (1997). Exploring the relationship between staff development and improvement in learning. *Journal for Staff Development*, 17(40), 34-38.
- Humphrey, J. W. (1988). *A longitudinal study of the consequences of full-day kindergarten through grade eight*. Evansville, IN: Evansville-Vanderburgh School Corporation.
- Iverson, S., & Tunmer, W. E. (1993). Phonological processing skills and the Reading Recovery program. *Journal of Educational Psychology*, 85, 112-126.
- Knight, S. L., & Stallings, J. A. (1995). The implementation of the Accelerated School model in an urban elementary school. In R. L. Allington & S. A. Walmsley (Eds.), *No quick fixes: Rethinking literacy programs in America's elementary schools* (pp. 236-251). New York: Teachers College Press.
- Madden, N. A., Slavin, R. E., Karweit, N. L., Donlan, L., & Waskik, B. A. (1991). *Success for All: Multi-year effects of a schoolwide elementary restructuring program* (Report No. 18). Baltimore: Center for Research on Effective Schools.
- Madden, N. A., Slavin, R. E., Karweit, N. L., Limermon, B. J., & Donlan, L. (1989). *Success for All: First-year effects of a comprehensive plan for reforming urban education* (Report No. 30). Baltimore: The Johns Hopkins University, Center for Research on Elementary and Middle Schools.
- Manset, G., St. John, E. P., Hu, S., & Gordon, D. (2002). Early literacy practices as predictors of reading related outcomes: Test scores, test passing rates, retention, and special education referral. *Exceptionality*, 10(1), 11-28.

- McCarthy, J., & Still, S. (1993). Hollibrook Elementary School. In J. Murphy & A. Hollinger (Eds.), *Learning from ongoing efforts* (pp. 63-83). Newberry Park, CA: Corwin.
- Miron, L. F., & St. John, E. P. (Eds.). (2003). *Reinterpreting urban school reforms: Have urban schools failed, or has the reform movement failed urban schools?* Albany: State University of New York Press.
- North Central Regional Educational Laboratory (NCREL). (2000). *Comprehensive school reform: Making good choices*. Oak Brook, IL: Author.
- Ohio State University, Reading Recover Project. (1998). *Literacy collaborative* (Information packet on Literacy Collaborative provided by program directors). Columbus, OH: Authors.
- Pinnell, G., Lyons, C. A., DeFord, D., Bryk, A., & Seltzer, M. (1994). Comparing instructional models for literacy education of high-risk first graders. *Reading Research Quarterly*, 29(1), 9-39.
- Ross, R. M., & Smith, L. J. (1994). Effects of Success for All model on kindergarten through second grade reading achievement teachers; adjustment and classroom-school climate at an inner city school. *Elementary School Journal*, 95, 121-138.
- Rowe, K. J. (1997). Factors affecting students' progress in reading: Key findings from a longitudinal study. In S. Swartz & A. Klein (Eds.), *Research in Reading Recovery* (pp. 53-101). Portsmouth, NH: Heinemann.
- Sheehan, R., Cryan, J. R., Wiechel, J., & Bandy, I. G. (1991). Factors contributing to success in elementary schools: Research findings for early childhood educators. *Journal of Research in Childhood Education*, 6, 66-75.
- Slavin, R. E. (2002). Evidence-based education policy: Transforming education practice and research. *Educational Researcher*, 31(7), 15-21.
- Slavin, R. E., & Fashola, O. S. (1998). *Show me the evidence! Proven and promising programs for America's schools*. Thousand Oaks, CA: Corwin Press.
- Snow, C., Burns, M., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- St. John, E. P., & Bardzell, J. S. (1999). *Improving early reading and literacy: A guide for developing research-based programs*. Bloomington: Indiana Education Policy Center.
- St. John, E. P., Griffith, A. I., & Allen-Haynes, L. (1997). *Families in schools: A chorus of voices in restructuring*. Portsmouth, NH: Heinemann.
- St. John, E. P., Manset, G., Chung, C. G., Simmons, A. B., & Musoba, G. D. (2000). *Research-based reading interventions: The impact of Indiana's Early Literacy Intervention Grant Program* (Policy Research Report #00-02). Bloomington: Indiana Education Policy Center.
- St. John, E. P., Manset, G., Chung, C. G., Simmons, A. B., Musoba, G. D., Manoil, K., et al. (2003). Research-based reading reform: The impact of state-funded interventions on educational outcomes in urban elementary schools. In L. F. Miron & E. P. St. John (Eds.), *Reinterpreting urban school reform: Have urban schools failed, or has the reform movement failed urban schools?* Albany: State University of New York Press.
- St. John, E. P., Ward, J. S., & Laine, S. W. M. (1999). *State policy on professional development: Rethinking the linkages to standard outcomes*. Oak Brook, IL: North Central Regional Educational Laboratory.
- Taylor, B. M., Anderson, R. C., Au, K. H., & Jaffy, E. R. (2000). Discretion in translation of research to policy: A case from beginning reading. *Educational Researcher*, 29(6), 16-26.
- Wong, K. K. (2003). Federal Title I as a reform strategy in urban schools. In L. F. Miron & E. P. St. John (Eds.), *Reinterpreting urban school reform: Have urban schools failed, or has the reform movement failed urban schools?* Albany: State University of New York Press.

Edward P. St. John is the Algo D. Henderson Collegiate Professor of Education in the Center for the Study of Higher and Postsecondary Education at the University of Michigan. Research interests include policy and practice in K-12 and higher education. His most recent publication is a 2004 volume he edited titled *Public Policy and College Access: Investigating the Federal and State Roles in Equalizing Postsecondary Opportunity*, Vol. 19: Readings on Equal Education (New York: AMS Press).

Genevieve Manset-Williamson is an associate professor in the School of Education at Indiana University. Research interests include literacy development and school reform as they pertain to students with learning disabilities. Her most recent publication is a 2002 article with S. Washburn titled "Administrators' Perspectives of the Impact of Mandatory Graduation Qualifying Examinations on Inclusive, Remedial, and Vocational Programming and Outcomes for Students With Learning Disabilities," in the *Journal of Special Education Leadership*, 15, 49-59.

Choong-Geun Chung is a researcher in the Smith Research Center at Indiana University. Research interests include statistical models for school reform, access and persistence in higher education, and issues in minority representation in special education. Her most recent publication is 2004 article with E. P. St. John, G. Musoba, and A. Simmons titled "Financial Access: The Impact of State Finance Strategies," in E. P. St. John (Ed.), *Public Policy and College Access: Investigating the Federal and State Roles in Equalizing Postsecondary Opportunity*, Vol. 19: Readings on Equal Education (New York: AMS Press).

Robert S. Michael is a research scientist in the Center for Evaluation and Education Policy at Indiana University. His research interests include school finance and research methodology.