DOCUMENT RESUME

ED 471 568 SP 041 281

AUTHOR	Petrosko, Joseph M.; Munoz, Marco A.
TITLE	A Collaborative Evaluation Model for Systematic Renewal of Teacher Education: Assessing the Effect of Professional Development Schools on Teachers and Students.
SPONS AGENCY	National Education Association, Washington, DC.
PUB DATE	2002-11-00
NOTE	31p.; Paper presented at the Annual Meeting of the American Evaluation Association (Washington, DC, November 6-10, 2002).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Achievement; College School Cooperation; Elementary Secondary Education; Higher Education; Partnerships in Education; *Preservice Teacher Education; *Professional Development Schools; *Program Evaluation; Teaching Methods; Urban Schools

ABSTRACT

This study described an evaluation model designed to evaluate the use and impact of Professional Development Schools (PDSs) in teacher education. It employed quantitative and qualitative data and was flexible enough to be used by practitioners in both education and training and development. Data were collected via observations, interviews, and surveys on teachers within PDS and non-PDS schools, examining such topics as instructional behaviors, explanations of instructional decisions, and attitudes and opinions about careers and the teaching profession. Data were also collected on student achievement, particularly in mathematics and science. Overall, effects of PDSs appeared to be classroom-specific and not school-wide. Observational data were relatively similar in both PDS and non-PDS schools. However, in five out of six variables selected, PDS classrooms had higher mean scores. The largest magnitude were on (1) classroom facility and classroom environment and (2) quality of instructional activity. Observational data also showed that teachers from PDS and non-PDS schools were generally similar in classroom teaching. Test scores indicated no statistically significant difference at the elementary level, but a significant difference at the high school level (though the magnitude of the difference was not large). (Contains 18 references.) (SM)

A Collaborative Evaluation Model for Systemic Renewal of Teacher Education: Assessing the Effect of Professional Development Schools on Teachers and Students

Joseph M. Petrosko
Nystrand Center of Excellence in Education
University of Louisville

Marco A. Munoz
Jefferson County Public Schools
Louisville, Kentucky

Paper Presented at the Meeting of the American Evaluation Association

November 2002

Note: The author's acknowledge the financial and technical support of the National Education Association (NEA) for the evaluation work described in this paper.

Abstract

This paper describes an evaluation model developed collaboratively by partners that are important to the success of teacher education: university-based researchers, school district researchers, and representatives of a teacher union. The model is being used to evaluate the use of Professional Development Schools (PDS) in teacher education. The model employs quantitative and qualitative data and is flexible enough to be used by practitioners in both education and training and development. Data from 10 schools analyzed to date reveal that effects of PDS may be classroom-specific and not school-wide. Implications for teacher education are discussed.

Keywords: Professional Development Schools; Preservice Teacher Education; College School Cooperation; Partnerships in Education; Teacher Improvement; Urban Schools.

A Collaborative Evaluation Model for Systemic Renewal of Teacher Education

The National Commission on Teaching and America's Future and multiple researchers (Ferguson & Womack, 1993; Darling-Hammond, 1997, 1999) have expressed the problematic condition of America's teaching force. For instance, in her review of state policy evidence on the relationship between teacher quality and student achievement, Linda Darling-Hammond (1999) argued that the most consistent highly significant predictor of student achievement in reading and mathematics is the proportion of well-qualified teachers in a state.

States are now focusing on teacher education as an important means for improving student achievement and are enacting legislation that raises standards for admission to teacher education and for certification. More recently, a report from the National Governors' Association Center for Best Practices urged states to consider establishing PDS partnerships as an effective model for teacher preparation programs and professional development for experienced teachers (Chronicle of Higher Education, 2000).

Schools and colleges of education have enacted programs such as Professional Development Schools (PDS) to address the need for field-based preparation. In the contemporary debate on the quality of teacher education, the use of the PDS model has emerged as a highly acclaimed model of teacher preparation (Book, 1996). However, there is a notable absence of studies attempting to make direct links between these innovations, the performance of teachers prepared in such programs, and the achievement of students in their classrooms. Indeed, Abdal-Haqq (1998) has documented the scarcity of information available to those interested in knowing more about the impact of PDS schools on teachers and K-12 students. While PDS partnerships have proliferated (Darling-Hammond, 1994; White, Rainer, Clift, & Benson, 1994), the

investment in these partnerships is based largely on theoretical assumptions about effective preparation programs of teachers. In the extant research on PDS partnerships, a notable absence of studies attempting to make direct links between PDS and the performance of teachers prepared in such programs and the achievement of students in PDS classrooms (Book, 1996; Bullough & Baughman, 1993; Ducharme & Ducharme, 1996; Zeichner, 1999).

This paper describes an evaluation model developed collaboratively by three partners that are important to the success of teacher education: university-based researchers, school district researchers, and representatives of a teacher union. The model was developed to serve the information needs of many stakeholders. The University of Louisville (U of L), located in a large urban area, is an institution with 12 colleges and approximately 22,000 students. The College of Education and Human Development has an extensive system of Professional Development Schools (PDS). In the process of conducting studies about PDS schools, U of L has developed a long-established research partnership with the nearby school district.

The Jefferson County Public Schools, situated in a state undergoing systemic school reform, is also facing the challenge of employing qualified teachers. The district is located in a large metropolitan area and has 150 schools serving approximately 93,000 students. The Jefferson County Teacher Association has the mission to serve their membership while promoting quality and equity in public education. As result, the teacher union is committed to the advancement of the teaching profession. The teacher union is one of the flagship locals within the NEA family of education associations.

The PDS are innovative institutions formed through partnerships between teacher education programs and K-12 schools. Their mission is professional preparation of candidates,

faculty development, inquiry directed at the improvement of practice, and enhanced student learning (NCATE, 2001). According to the National Council for Accreditation of Teacher Education, PDSs have distinct characteristics.

> They are learning environments that support candidate and faculty development within the context of meeting all children's needs. PDS partners are guided by a common vision of teaching and learning, which is grounded in research and practitioner knowledge. PDS partners share responsibility for professionals and students; they blend their expertise and resources to meet their shared goals. (NCATE, 2001, p. 2)

PDS partners work together over time, building relationships and commitment to their shared goals. They develop new strategies, roles, and relationships to support their work. Together, they move to institutionalize their partnerships so that it is supported and becomes part of their institutions' expectations. At the most advanced stages of development, PDS partnerships influence policies and practices at the district and state levels.

As implemented by the University of Louisville, PDS involves teacher education students spending substantial amounts of time at a public school (i.e., the PDS) taking university classes, assisting experienced teachers, and instructing students (Book, 1996; Holmes Group, 1995). This approach to teacher education contrasts with the traditional approach in which aspiring teachers take almost all of their classes on campus, and then experience student teaching for a relatively brief period.

Evaluation of PDS has been very limited, both in Louisville or everywhere else the approach has been implemented. With the impetus of a grant from the National Education Association (NEA), an evaluation was designed for the Louisville PDS-based system of teacher

education. The major purpose of the collaborative model was to gain insight into the impact of professional development schools on: (a) teacher behavior, attitudes, and opinions, and (b) student achievement, particularly in mathematics and science. The overall goal of the evaluation is systemic and data-based renewal of teacher education. While this collaborative evaluation model emerged from the field of education, it has features that could be adapted by many practitioners, especially those with a training and development focus, including human resource development (HRD) professionals.

The presence of the PDS sites allowed the possibility of conducting research on this particular type of approach to teacher preparation about which a wealth of information is available. The literature on PDS research also helped to frame this study. According to Valli, Cooper, and Frankes (1997), there is limited research on the relationship between PDS and improved classroom practice. Research on student learning focuses more on teaching inputs and less on learning outcomes.

Ross, Brownell, Sindelar, and Vandiver (2000) argued that researchers are slow to explore the relationships to student achievement because they are skeptical about the adequacy of achievement tests to measure PDS outcomes. The authors concluded their review of research on PDS mentioning that there are a number of action research studies in the literature, but that these do not lend themselves to generalization. In agreement with Valli et al. in recognizing the paucity of systematic studies, Ross et al. noted "the research we most need from current PDS efforts is research that documents explicitly the positive impact of PDS efforts on student learning" (p. 17).

Evaluation Model

Key features of the evaluation model developed for the PDS project are shown in Table 1. Shown in Column 1 are major evaluation components. Columns two and three describe the two major arms of the evaluation: (a) teacher outcomes, and (b) student achievement. This model was developed for the relatively common situation of a program having *proximate* outcomes and *distal* outcomes (Rossi, Freeman, & Lipsey, 1999).

PDS has proximate outcomes. These are measured on the teachers impacted by the PDS--in this case teachers working in a PDS, as contrasted with teachers working at other schools in the district without the PDS designation. In the literature on PDS, certain features of the school are hypothesized to impact professional personnel (Book, 1996). Schools are expected to be open to innovation, supportive of teacher professional development, and likely to be a more positive working environment for educational reform. Presumably, differences could be apparent on these dimensions if a PDS school were compared with a school that was not PDS, but was in other ways similar (e.g. in demographic characteristics of students).

PDS has distal outcomes. It is entirely appropriate that the proximate outcomes of PDS be improved organizational climate, a positive work environment for teachers, and instructional progressivism. However, the fundamental purpose of school remains to produce positive student outcomes.

A meaningful distal outcome of any approach to teacher education (e.g., PDS) is student achievement. Does the achievement of K-12 students in PDS differ from that of students in non-PDS? We viewed this question as more exploratory than confirmatory. Authors who have described the benefits of PDS have rarely discussed student educational outcomes. Improvement

in such outcomes is not explicitly ruled out, but neither are strong claims made for their existence. It is an open question whether such outcomes can be verified.

According to Chen and Rossi (1983), the use of theoretical models in program impact assessment can heighten the power of experimental designs and compensate for some deficiencies in the quasi-experimental designs. In the language of Rossi, Freeman, and Lipsey (1999) the distal outcome of student achievement reflects a conceptual hypothesis of the implicit program impact theory of PDS.

Table 1

Evaluation Model for Professional Development Schools (PDS) as an Approach to Teacher Education

Evaluation Component	Evaluation Focus I. Proximate Outcomes: Teacher Behaviors, Attitudes, and Opinions	Evaluation Focus II. Distal Outcomes: Student Achievement
Purpose of the evaluation	Providing potentially useful information for decision-making about PDS in the context of a teacher education program (formative)	Providing exploratory evaluation/ research data on ultimate impact of PDS
General evaluation question	Comparing teachers employed in PDS schools and non-PDS schools, are there differences in teacher instructional behaviors, and teacher attitudes?	Are there differences in achievement of K-12 students in PDS schools and non-PDS schools?
Data type	Quantitative and Qualitative	Quantitative
Data gathering technique	School visitation of individual teachers	Retrieval of archival data from school district student record system
Categories of variables	Data on teachers within PDS and non-PDS schools: Instructional behaviors, explanations of instructional decisions, attitudes and opinions about career and the teaching profession	Data on individual students within PDS and non-PDS schools: Demographic data, achievement scores, attendance
Analysis methods	Descriptive statistics (for observational and survey data) Narrative summaries and thematic analysis (for interviews and observations)	Descriptive and inferential statistics (*t*-tests, ordinary least squares regression, multi-level analysis/hierarchical linear modeling)
Reporting methods	Narrative and statistical summaries	Statistical summaries

Application of the Evaluation Model

The following two sections provide some details of how the evaluation model was applied and some findings obtained in the first year of the model's application. We divide the discussion into two major sections, one for each of the outcomes of the PDS, proximal and distal.

Evaluation Focus I. Proximate Outcomes: Teacher Behaviors, Attitudes, and Opinions

Purpose of the evaluation

The purpose of Evaluation Focus I was to provide potentially useful information for decision-making about PDS. This was viewed as formative rather summative decision-making, since PDS is an established and on-going teacher education program. Data on teacher behaviors in PDS and non-PDS environments are useful to teacher educators, since such data bear on both the objectives of university instruction and the goals of PDS. For example, if PDS is to encourage modern approaches to instruction, this should be evident in teacher classroom activities.

General evaluation question

At its most general, the central evaluation question is: Comparing teachers employed in PDS schools and non-PDS schools, are there differences in teacher instructional behaviors, and teacher attitudes? Teachers in a PDS school should be equal to or greater than the level of progressive instructional practices compared to teachers in a non-PDS school.

Data type and data gathering technique

The types of data obtained from teachers were both quantitative (e.g., coded observation forms) and qualitative (e.g., open-ended interview questions). Data were gathered by visiting individual teachers in their classrooms.

Categories of variables and analysis methods

There were several categories of data collected on teachers within both PDS and non-PDS schools. These included: (a) Instructional behaviors (from observation form), (b) explanations of instructional decisions (interview), and (c) attitudes and opinions about careers and the teaching profession (questionnaire). Analysis methods for quantitative data included compiling descriptive statistics for observational and survey data. Narrative summaries and thematic analysis were produced for interviews and observations. This report will contain a selected set of the quantitative results.

Observational data on teachers

We collected classroom observation data from a total of 16 teachers. Eight came from PDS schools (4 from an elementary school and 4 from a high school) and eight came from non-PDS schools (4 from an elementary school and 4 from a high school). Observation instruments were local adaptations of instruments designed by the Appalachian Rural Systemic Initiative (ARSI). ARSI is a curriculum improvement consortium for K-12 mathematics and science instruction, and is supported by the National Science Foundation (NSF). We acknowledge the helpful cooperation of Dr. Ron Pelfrey, who granted permission for adaptations.

Classroom observation instruments required raters to observe and rate aspects of the classroom and of student-teacher interactions during a mathematics or science lesson. The five-page forms required numerous ratings and qualitative comments. This report provides a brief summary of overall ratings in several areas that are important to the instructional experience.

Classroom observations of the 16 teachers in the study occurred in April and May 2002. Each teacher was observed twice—once by a University of Louisville researcher and once by an experienced observer employed for the project. Brief explanations of each rating scale item are as follows. In all cases, the higher the number for a rating (i.e., close to 5 or 7) the more the classroom was rated as engaging in practices consistent with the goals of PDS.

Rating 1 --Physical Setting /Classroom Environment This measured whether the student seating arrangement, the classroom facility and the classroom environment were conducive to active instruction.

Ratings 2 and 3--Instruction Lessons were rated (1 to 5) in terms of whether they involved instructional activity that was engaging to students. Areas that contributed to this rating included how well the curriculum materials were used, amounts of teacher and student led interaction, and amount of hands-on activity. In addition, each lesson was rated (1 to 7) as to the appropriateness of instructional activity: how much it fit the purported goals of the lesson.

Ratings 4 and 5—Classroom Questioning Lessons were rated on the classroom questioning of teachers. The variable called Questioning (1 to 5) was a rating of the teacher in terms of: (a) using an effective balance of convergent and divergent questions, and (b) encouraging questions by students. The variable Appropriateness of Questioning Strategies (1 to 7) was a rating of the teacher in terms of how well the questioning strategy fit the apparent goals of the lesson.

Ratings 6—Higher Order Thinking Lessons were rated (1 to 5) on how effectively the teacher used instruction to encourage higher-order skills in students.

Table 2 shows the percentage of agreement between the two observers per classroom. When exact agreement was calculated (e.g., both raters gave a 3, or both raters gave a 4) percentages were

not high. They ranged from 13% to 62%. However, when the criteria were liberalized, the percentages improved. Percentages of agreement within one point were calculated (e.g., one rater gave a 3, the other a 2, 3, or 4). These ranged from 73% to 100%.

Table 2

Percentage of Agreement Between Two Classroom Observers For Six Observation Criteria

Observation Criteria	Percentage of Exact Agreement	Percentage of Agreement within one point
1. Physical Setting /Classroom Environment	13	87
2. Lesson/Instructional Activity	46	100
3. Appropriateness of Instruction	40	100
4. Questioning	53	87
5. Appropriateness of Questioning	62	92
6. Instruction for Higher Order Skills	53	73

We decided data were reliable enough to use, at least for evaluation/research purposes. Ratings were averaged across the two observations per teacher. Thus, the data that are reported here represent the mean of two independent judgments of teacher performance.

Table 3 shows mean scores on six classroom observation criteria. Each covered an important dimension of classroom quality that could plausibly be different when comparing PDS and non-PDS schools. Data were relatively similar in both PDS and non-PDS classrooms. However, in five out of six of the variables selected, PDS classrooms had higher mean scores. The largest magnitude of difference exited for ratings 1 and 2. PDS exceeded non-PDS classrooms on: (a) the classroom facility and the classroom environment, and (b) the quality of the instructional activity. Figures 1

and 2 show means on these two rating criteria, and also the 95% confidence intervals for PDS and non-PDS schools.

Table 3

Average Classroom Observations for Teachers from Professional Development Schools (PDS) and Non-PDS Schools

Classroom Observation Ratings on 8 PDS classrooms and 8 non-PDS classrooms

		Mean	Std. Deviation	N
1. Physical Setting/Classroom Environment (1-5)	PDS	4.3750	.44320	8
	non-PDS	3.5625	.32043	8
	Total	3.9688	.56181	16
2. Lesson/Instructional Activity (1-5)	PDS	4.3125	.59387	8
	non-PDS	3.8125	.25877	8
	Total	4.0625	.51235	16
3. Appropriateness of Instruction (1-7)	PDS	5.9375	.86344	8
	non-PDS	5.5000	.53452	8
	Total	5.7188	.72958	16
4. Questioning (1-5)	PDS	3.8750	.69437	8
	non-PDS	3.4375	.41726	8
	Total	3.6563	.59774	16
5. Appropriateness of Questioning Strategies (1-7)	PDS	5.4375	.90386	8
	non-PDS	5.5625	.41726	8
	Total	5.5000	.68313	16
6. Instruction Leading to Higher Order Skills (1-5)	PDS	3.8750	.95431	8
	non-PDS	3.5625	.49552	8
	Total	3.7187	.75208	16

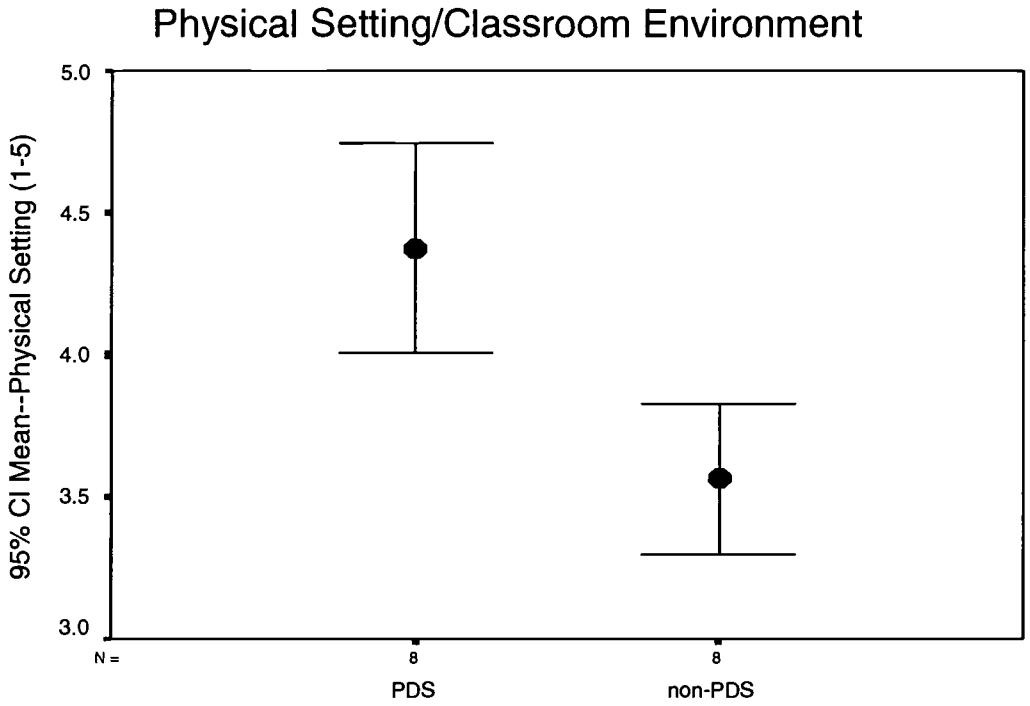

Figure 1. Average ratings on Physical Setting/Classroom Environment.

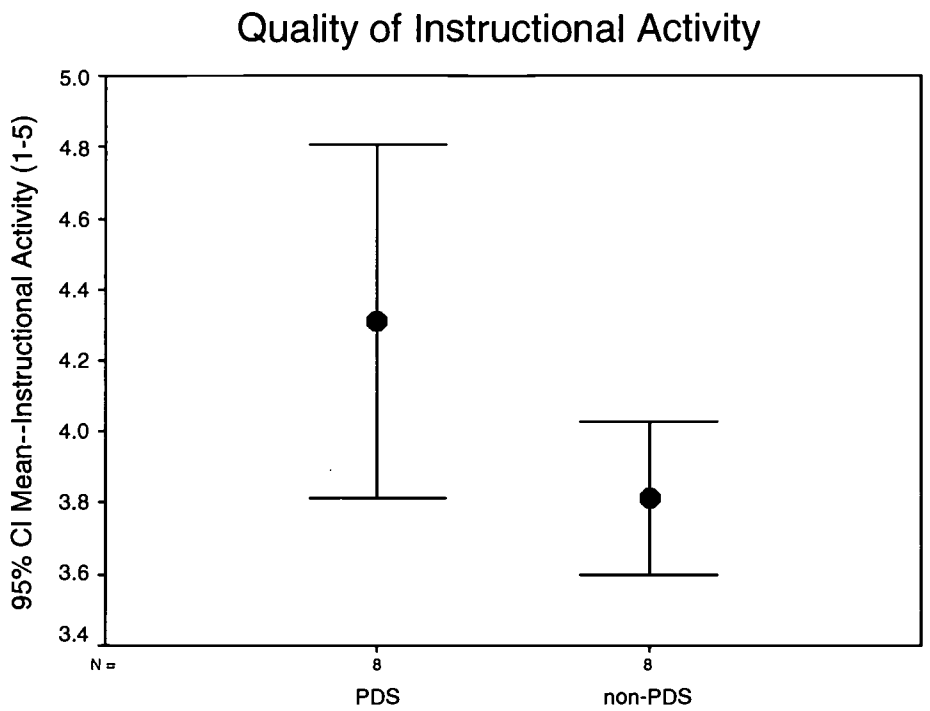

Figure 2. Average ratings on Quality of Instructional Activity.

Summary of observational data

Teachers from PDS schools and non-PDS schools were generally similar in classroom teaching. When differences were found, teachers in PDS had higher scores. Teachers were interviewed in conjunction with the observation process. Responses to open response items on the interview form (qualitative data) are not described in this report.

Responses to teachers to selected items from a questionnaire

We administered a questionnaire to teachers who participated in the classroom observation phase of the study. A subset of items was selected for analysis in this report. All deal with the general construct of school culture—organizational climate of the school affecting teacher work and professional development. Items dealt with whether teachers perceived a common vision in the school, whether the school culture encouraged professional development, and whether the school culture encouraged teaching for higher order understanding.

Analysis revealed that PDS mean scores exceeded those of non-PDS classrooms (see Table 4). Figure 3 shows mean scores for the items, with both PDS and non-PDS teacher data combined. Averages were high for the combined means on all three items.

Table 4

Average Ratings on Three Aspects of School Culture for Teachers from Professional Development Schools (PDS) and Non-PDS Schools

Ratings of Three Aspects of School Culture for Teachers in PDS and not in PDS

		Mean	Std. Deviation	N
SC1 Faculty at this school share a common vision regarding goals. (1-5)	PDS	4.6250	.5175	8
	non-PDS	4.0000	.9258	8
	Total	4.3125	.7932	16
SC6 The school culture encourages faculty to continue their professional development. (1-5)	PDS	4.6250	.7440	8
	non-PDS	4.2500	1.0351	8
	Total	4.4375	.8921	16
SC9 The school culture encourages faculty to teach for understanding and higher order reasoning. (1-5)	PDS	4.7500	.4629	8
	non-PDS	4.5000	.5345	8
	Total	4.6250	.5000	16

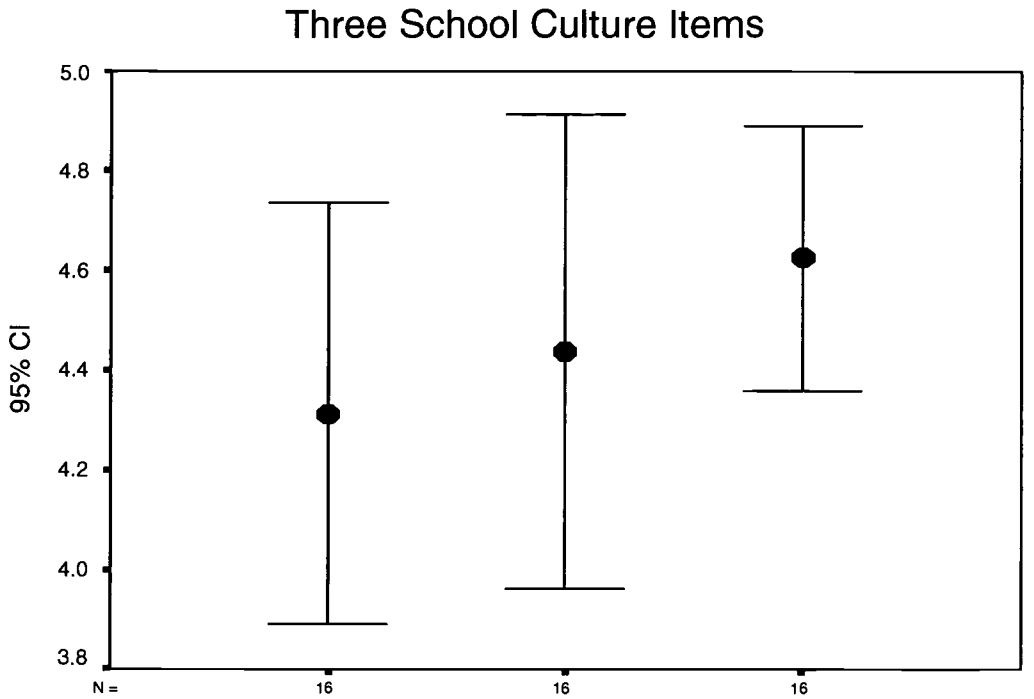

Figure 3. Average ratings (PDS and non-PDS combined) on three School Culture ratings: 1. Faculty in the school have common vision, 2. School culture encourages professional development, 3. School culture encourages teaching for understanding

Evaluation Focus II. Distal Outcomes: Student Achievement

Purpose of the evaluation

The purpose of Evaluation Focus II was to provide exploratory evaluation/research data on the ultimate impact of PDS. The research literature is sparse on the effects of PDS on student achievement (in contrast to a substantial literature on its effects on adults). This evaluation focus aims to provide overall enlightenment and not definitive evidence for decision-making.

General evaluation question

The question for Evaluation Focus II was: Are there differences in achievement of K-12 students in PDS schools and non-PDS schools? The most hopeful outcome would be that achievement of PDS school students would be equal to or greater than that of students in a non-PDS school. However, it must be borne in mind that schools are selected to be PDS partly because they have needs that are seen as being addressed by the university affiliation inherent in the PDS designation. PDS sites are often selected because the school serves at risk students (at least in part) or because teachers in the school have professional development needs. This complicates evaluations, since achievement data may be systematically biased against PDS prior to any comparisons with other schools.

Data type and data gathering technique

Data were quantitative, primarily test score data from students. Achievement scores and other student information were retrieved as archival data from school district student record system.

Categories of variables and analysis methods

There were several categories of data collected on students within both PDS and non-PDS schools. These included: (a) demographic data, (b) achievement scores, and (c) student attendance data. Quantitative analysis techniques included calculation of: (a) descriptive statistics, (b) standard types of inferential statistics (*t*-tests, ordinary least squares regression), and (c) newly emerging inferential statistics (multi-level analysis/hierarchical linear modeling).

Data came from four elementary schools and six high schools. Table 5 below shows the key information about these schools. At both the elementary and high school levels, data were obtained from PDS and matched non-PDS schools. The non-PDS control schools were selected by the JCPS Office of Research, Planning, and Accountability. Every school in the district has a match—a school that it can be paired with for comparison purposes during research studies or program evaluations. Matches are on the basis of student demographic variables and average achievement levels.

Table 5

Student Outcomes Data from Schools in the Study

School Type	Number of individual student test scores
Elementary 2 PDS and 2 matched control schools	236
High School 3 PDS and 3 matched control schools	1374

Analysis of elementary school test score data

Analysis consisted of a variety of descriptive and inferential statistics, comparing achievement for PDS schools and non-PDS schools. Only some results are presented in this report. Table 6 shows outcomes of the major analyses that were pursued. No statistically significant

differences were found between PDS schools and matched control schools. Comparisons included nationally standardized tests (CTBS) and the state of Kentucky assessment (KIRIS). Analysis also revealed no differences between PDS and non-PDS school students in percentages of days attending school.

Table 6

Main Results of Elementary Student Outcomes Data

Outcome Variable	Elementary School Data Analysis
CTBS	No significant difference between PDS and non-PDS on Grade 3 CTBS scores.
KIRIS science KIRIS math	No significant difference between PDS and non-PDS on Grade 4 KIRIS science scores and Grade 5 KIRIS math scores.
KIRIS science (Controlling for previous achievement, sex, ethnicity, and free and reduced lunch status)	Dependent variable was Grade 4 KIRIS science scores. No significant increase in variance by using the variable PDS/ non-PDS as a predictor, after controlling for grade 3 CTBS scores, sex, ethnicity, and free and reduced lunch status.
KIRIS math (Controlling for previous achievement, sex, ethnicity, and free and reduced lunch status)	Dependent variable was Grade 5 KIRIS math scores. No significant increase in variance by using the variable PDS/ non-PDS as a predictor, after controlling for grade 3 CTBS scores, sex, ethnicity, and free and reduced lunch status.
CTBS (In year after student left elementary school)	No significant difference between PDS and non-PDS on Grade 6 CTBS scores.
Student Attendance	No significant difference between PDS and non-PDS on percentage of days students attended school 4 separate years.

Note. CTBS = Comprehensive Tests of Basic Skills (Nationally standardized test)
KIRIS = Kentucky Instructional Results Information System (State assessment)

Table 7 shows the third analysis described above. The dependent variable in a multiple regression analysis was Grade 4 KIRIS science scores. Hierarchical entry of variables was used. In the last regression model it was found there was no significant increase in variance by using the variable PDS/ non-PDS as a predictor, after controlling for grade 3 CTBS scores, sex, ethnicity, and free and reduced lunch status. A similar finding occurred when the same predictor variables were used with Grade 5 KIRIS mathematics scores.

Table 7

Performance for Students in Four Elementary Schools: Two Professional Development Schools (PDS) And Two Matched Non-PDS Schools. Four Regression Models Predicting KIRIS Grade 4 Science Scores

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.486[a]	.236	.233	.5238	.236	66.251	1	214	.000
2	.569[b]	.324	.314	.4953	.087	13.661	2	212	.000
3	.590[c]	.348	.330	.4896	.025	2.659	3	209	.049
4	.593[d]	.351	.329	.4897	.003	.889	1	208	.347

a. Predictors: (Constant),CTBS math grd. 3

b. Predictors: (Constant), CTBS math grd. 3, CTBS language grd. 3, CTBS reading grd. 3

c. Predictors: (Constant), CTBS math grd. 3, CTBS language grd. 3, CTBS reading grd. 3, SEX3, LUNCH2, RACE3

d. Predictors: (Constant), CTBS math grd. 3, CTBS97 language grd. 3, CTBS reading grd. 3, SEX3, LUNCH2, RACE3, PDS3

Note. Last model indicates that a statistically non-significant amount of variance is predicted by school: PDS school students were not different than non-PDS, controlling for previous achievement in mathematics, language, reading, ethnicity of student, sex of student, and students' free and reduced lunch status.

Analysis of high school test score data

Similar to the elementary school data, analysis of high schools consisted of a variety of descriptive and inferential statistics, comparing achievement for PDS schools and non-PDS schools. Table 8 shows outcomes of the major analyses. On the state standardized achievement test (CATS) statistically significant differences were found between PDS and matched control schools, favoring PDS schools. However, the magnitude of these relationships was not large. Analysis also revealed no differences between PDS and non-PDS school students in percentages of days attending school.

Table 8

Main Results of High School Student Outcomes Data

Outcome Variable	High School Test Score Data Analysis
CTBS	PDS higher than non-PDS on Grade 9 CTBS scores
CTBS math (Controlling for sex, ethnicity, and free/reduced lunch)	Dependent variable was Grade 9 CTBS math scores. Small (.027 proportion of variance), but statistically significant increase in variance by using the variable PDS/non-PDS as a predictor, after controlling for, sex, ethnicity, and free and reduced lunch status. PDS has a positive effect.
CTBS Total (Controlling for sex, ethnicity, and free/reduced lunch)	Dependent variable was Grade 9 CTBS Total scores. Small (.024 proportion of variance), but statistically significant increase in variance by using the variable PDS/ non-PDS as a predictor, after controlling for sex, ethnicity, and free and reduced lunch status. PDS has a positive effect.
CATS science	PDS higher than non-PDS on Grade 11 CATS science scores and Grade 11 CATS math scores.
CATS science (Controlling for sex, ethnicity, and free and reduced lunch status)	Dependent variable was Grade 11 CATS science scores. Very small (< .01 proportion of variance), but statistically significant increase in variance by using the variable PDS/non-PDS as a predictor, after controlling for, sex, ethnicity, and free/reduced lunch. PDS has a positive effect.
CATS math (Controlling for sex, ethnicity, and free and reduced lunch status)	Dependent variable was Grade 11 CATS math scores. Very small (< .01 proportion of variance), but statistically significant increase in variance by using the variable PDS/non-PDS as a predictor, after controlling for, sex, ethnicity, and free/reduced lunch. PDS has a positive effect.
Student attendance	No significant difference between PDS and non-PDS on percentage of days students attended school over 4 separate years

Table 9 shows the fifth analysis described above. The dependent variable in a multiple regression analysis was Grade 11 CATS science scores. Hierarchical entry of variables was used. In the last regression model it was found there was a small, but statistically significant increase in variance by using the variable PDS/ non-PDS as a predictor, after controlling for CTBS scores, sex, ethnicity, and free and reduced lunch status. A similar finding occurred when the same predictor variables were used with Grade 11 CATS mathematics scores.

Table 9

Performance for Students in Six High Schools:
Three Professional Development Schools (PDS) And Three Matched Non-PDS Schools
Four Regression Models Predicting CATS Grade 11 Science Scores

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.631[a]	.399	.398	.9078	.399	872.142	1	1315	.000
2	.679[b]	.461	.459	.8604	.062	75.349	2	1313	.000
3	.693[c]	.480	.478	.8458	.019	16.283	3	1310	.000
4	.694[d]	.482	.480	.8443	.002	5.621	1	1309	.018

a. Predictors: (Constant), CTBS math grd. 9

b. Predictors: (Constant), CTBS math grd. 9, CTBS language grd. 9, CTBS reading grd. 9

c. Predictors: (Constant), CTBS math grd. 9, CTBS language grd. 9, CTBS reading grd. 9, RACE3, SEX3, LUNCH2

d. Predictors: (Constant), CTBS math grd. 9, CTBS language grd. 9, CTBS reading grd. 9, RACE3, SEX3, LUNCH2, PDS3

Note. Last model indicates that a statistically significant (but small) amount of variance is predicted by school: PDS school students were superior to non-PDS, controlling for previous achievement in mathematics, language, reading, ethnicity of student, sex of student, and students' free and reduced lunch status.

Summary of comparisons of student test score data from PDS schools and non-PDS schools

Being a student in a PDS school confers neither a great advantage nor a disadvantage in terms of achievement. At the elementary school level PDS schools were not statistically different than non-PDS schools. At the high school level, PDS school students had higher achievement scores than non-PDS students. However the magnitudes of the differences were not large, especially after controlling for background variables.

Finally, an exploratory exercise was conducted using Hierarchical Linear Modeling (HLM). HLM allows the researcher to accurately estimate a school effect on a dependent variable measured on students. Here, the individual student test score data are defined as the dependent variable of a Level 1 model. The Level 2 model has the coefficients of the Level 1 model as its dependent variables.

The authors calculated several HLM equations with the high school data set, using Grade 11 CATS mathematics scores as the Level 1 dependent variable. Tables 10 and 11 show the results. The last model shown in Table 10 has its key results displayed in Table 11. PDS was not a significant predictor of CATS Math scores, a result consistent with simpler statistical analyses. Additional work needs to be done with HLM and PDS schools. This type of analysis works best when the number of schools is large, or alternatively, the level of analysis is changed so that PDS versus non-PDS classrooms are used rather than PDS (whole schools) versus non-PDS (whole schools).

Table 10
Summary of Several Hierarchical Linear Models Using
Grade 11 CATS Math as the Dependent Variable

HLM Model	Key results
One way ANOVA	There is significant variance among school means for Grade 11 CATS math. This justifies further HLM analysis.
Means as outcomes, one Level 2 predictor	There is significant relationship between Level 2 predictor CTBS Math School Average and school means on the outcome variable, Grade 11 CATS math
Intercept and slopes as outcomes model. CTBS Math is Level 1 Predictor, PDS status (1 or 0) and school average CTBS Math are Level 2 Predictors.	The average school mean on outcome variable Grade 11 CATS Math is significantly different than zero. Of the two Level 2 predictors, PDS is not significant and CTBS Math is significant. *Interpretation: PDS status does not significantly predict school averages in Grade 11 CATS Math. Result is consistent with ordinary least squares multiple regression.*

Table 11
Key Result of Intercept and Slopes as Outcomes HLM Model Using Grade 11 CATS Math as the Dependent Variable and 3 predictors, One Level 1 and Two Level 2

```
The outcome variable is GRADE 11 CATS MATH
Level 1 predictor is CTBSstudent (measured at the individual student level)
Level 2 predictors are:
                      PDS (1=yes, 0=no)
                      CTBSMATH (CTBS average for the school)

 Final estimation of fixed effects:
 ----------------------------------------------------------------------------
                                      Standard              Approx.
    Fixed Effect          Coefficient   Error      T-ratio   d.f.     P-value
 ----------------------------------------------------------------------------
 For       INTRCPT1, B0
    INTRCPT2, G00           4.782553   0.085156    56.162        3     0.000
         PDS, G01           0.353338   0.129107     2.737        3     0.068
    CTBSMATH, G02           0.074404   0.008770     8.484        3     0.000
 For  CTBSstudent slope, B1
    INTRCPT2, G10           0.074656   0.003242    23.030        3     0.000
         PDS, G11          -0.000932   0.004912    -0.190        3     0.862
    CTBSMATH, G12          -0.000352   0.000340    -1.035        3     0.377
 ----------------------------------------------------------------------------
```

Discussion

The study described an evaluation model developed collaboratively by partners in teacher education, namely (a) university-based researchers, (b) school district researchers, and (c) representatives of a teacher union. The collaborative model was used to evaluate the use and impact of PDS schools in teacher education. PDS schools involve teacher education students spending substantial amounts of time at a public school (i.e., PDS school site) to take their university classes and instructing students. This approach to teacher education contrasts with the traditional approach in which aspiring teachers take almost all of their classes on campus.

The evaluation model employed a mixed model approach by combining quantitative and qualitative data to gain insight into the impact of PDS schools on (a) teacher behavior, attitudes, and opinions, and (b) student achievement, particularly in mathematics and science. This model was developed for assessing program proximate outcomes and distal outcomes (Rossi, Freeman, & Lipsey, 1999). The distal outcome of student achievement reflects a conceptual hypothesis of the implicit program impact theory of PDS.

Observational data were relatively similar in both PDS and non-PDS schools. However, in five out of six of the variables selected, PDS classrooms had higher mean scores. The largest magnitude were on (a) the classroom facility and the classroom environment and (b) the quality of the instructional activity. Furthermore, observational data showed that teachers from PDS and non-PDS schools were generally similar in classroom teaching. When differences were found, however, teachers in PDS schools had higher scores.

Test scores data analyses indicated no statistically significant difference at the elementary school level, but a significant one at the high school level. PDS high schools had higher grade 9

CTBS scores, higher grade 11 CATS science scores, and higher grade 11 CATS math scores; however, the magnitude of the differences was not large, especially after controlling for background variables. The HLM analysis showed that PDS status does not significantly predict school averages in grade 11 CATS mathematics. In this sense, the result was consistent with the Ordinary Least Squares (OLS) multiple regression (Tabachnick & Fidell, 1996). Student attendance did not show statistically significant difference at any school level.

This study, as any research investigation, has multiple limitations. Further research is needed to address these limitations. At least two kinds of additional analyses should be pursued. First, the analyses reported above involved data from entire schools. Additional analyses would be beneficial, especially with student data from teachers *most involved* with PDS compared with teachers that are *not involved* with PDS. Secondly, thorough application of hierarchical linear modeling (HLM) needs to be performed with achievement data. For example, classroom level variables might prove very valuable to analyze.

In concluding this study, it is important to encourage educational researchers to continue studying the impact of teachers on student learning and achievement. Teachers are the most important variable in student learning. Policy makers need enlightenment about the particularities (i.e., patterns of behavior and practice) of effective teachers in PDS and non-PDS schools. Teacher educators need to know what how to prepare effective teachers or help those who are less effective become more effective. Dissemination of the findings on this kind of project will help policy makers, teacher preparation programs, school systems, and other major stakeholders in a high-stakes testing environment.

References

Abdal-Haqq, I. (1998). *Professional Development Schools: Weighting the evidence.* Thousand Oaks, CA: Corwin Press.

Book, C. L. (1996). Professional development schools. In J. Sikula, T. Buttery, & E. Guyton (Eds.). *Handbook of research on teacher education* (2nd ed.) (pp. 194-210). New York: Macmillan.

Bullough, R. V., & Baughman, K. (1993). Continuity and change in teacher development: First year teacher after five years. *Journal of Teacher Education, 44(2),* 86-95.

Chen, H. T. & Rossi, P. H. (1983). Evaluating with sense: A theory-driven approach. *Evaluation Review, 7(3),* 283-302.

Darling-Hammond, L. (1994). *Professional Development Schools: Schools for developing a profession.* New York: Teachers College Press.

Darling-Hammond, L. (1997). *Doing what matters most: Investing in quality teaching.* New York: National Commission on Teaching and America's future.

Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence.* Seattle: Center for the Study of Teaching and Policy, University of Washington.

Ducharme, E. R., & Ducharme, M. K. (1996). Needed research in teacher education. In Sikula, J. Buttery, J. J., and Guyton, E. (Eds.)., *Handbook of research on teacher education* (2nd ed.). New York: MacMillan.

Ferguson, P., & Womack, S. T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education, 44(1),* 55-63.

Holmes Group. (1995). *Tomorrow's schools of education.* East Lansing, MI: Author.

NCATE (2001). *Standards for Professional Development Schools.* Washington, DC: Author.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Ross, D., Brownell, M., Sindelar, P., & Vandiver, F. (February 2000). Research from professional development schools: Can we live up to the potential? *Peabody Journal of Education.*

Rossi, P. H., Freeman, H.E., & Lipsey, M. W. (1999). *Evaluation: A systematic approach* (6th ed.). Thousand Oaks, CA: Sage.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.

The Chronicle of Higher Education (26 January, 2000). *Governors urge help from states on university-school teaching partnerships.*

Valli, L., Cooper, D., & Frankes, L. (1997). Professional development schools and equity: A critical analysis of rhetoric and research. In M. Apple (Ed.), *Review of Research in Education, Volume 22.* Washington, DC: American Educational Research Association.

White, A., Reiner, G., Clift, R., & Benson, S. (1994). School/University partners in research: The potential and the problems. In M. J. O'Hair and S. J. Odell (Eds.)., *Partnership in Education: Teacher education Yearbook II* (pp. 27-44). Fort Worth, TX: Harcourt Brace.

Zeichner, K. (1999). The new scholarship in teacher education. *Educational Researcher 28(9),* 4-15.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: A Collaborative Evaluation Model for Systemic Renewal of Teacher Education	
Author(s): Joseph M. Petrosko and Marco A. Munoz	
Corporate Source: University of Louisville / JCPS	Publication Date: 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ☑	**Level 2A** ☐	**Level 2B** ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Marco Munoz	Printed Name/Position/Title: Marco A. Munoz Evaluation Specialist	
Organization/Address: 3332 Newburg Rd Louisville, KY 40218	Telephone:	FAX:
	E-Mail Address:	Date: 12/10/02

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: info@ericfac.piccard.csc.com
WWW: http://ericfacility.org